DECENTRING DEVICES:

Developing Quali-Quantitative Techniques for Studying Controversies with Online Platforms

David Moats

Sociology, Goldsmiths, University of London
Thesis submitted for the degree of Doctor of Philosophy (PhD) 2015

DECLARATION

I hereby declare that the work presented in this thesis is my own. Wherever contributions of others are involved, these are clearly acknowledged.

David Moats

ACKNOWLEDGMENTS

As STS scholars are well aware, any research project is always a *distributed accomplishment,* but this is more obviously the case in the Wild West of digital research. Firstly, I am grateful to the ESRC for funding such an experimental project and of course to my brilliant supervisors Noortje Marres and David Oswell who set the bar very high and continued to make insightful and challenging comments throughout. I could not have asked for a better pair of voices jostling in my head. A different version of the argument in Chapter II and the tool developed in Chapter VI have appeared in the article *Mapping Controversies with Social Media: The Case for Symmetry* with Noortje Marres (2015). I also want to thank the entire team at the Digital Methods Initiative and the Média Lab for their support, both technical and intellectual. I have benefitted considerably from discussions with Richard Rogers, Tommaso Venturini, Bernhard Rieder, Sabine Niederer and Carolin Gerlitz. A collaboration with Erik Borra led to the creation of the URL sequencer in Chapter VI, which he was kind enough to incorporate into the DMI TCAT interface. Gabriele Colombo, Federica Bardelli, Carlo di Gaetano and Alessandro Brunetti from Density Design all made helpful suggestions on the visualisations presented. I have enjoyed more casual but no less profound chats with past and present members of CSISP at Goldsmiths: Michael Guggenheim, Joe Deville, Martin Savransky, Anna-Maria Teohari-Herman, Jess Perriam, Sveta Milyaeva, Vera Ehrenstein, Manuel Tironi, Andreas Birikbak, Irina Papazou and especially my excellent upgrade examiners Mike Michael and Alex Wilkie and office-mate Laurie Waller. I am also indebted to Nick Couldry, Vic Seidler, Anna Skarpelis and everyone at the NYLON reading group for helping me stay connected to my qualitative, sociological roots; Don Slater for setting me down my current path; Liz Moor for being so generous with her support and advice and Les Back for his passion for sociology. Also thanks go to my fellow travellers Miranda Iossofidis, Jim Kaufman, Anna Bull, Carolina Ramirez, Alice Corble, Felipe Palma, Francisco Calafate, Phil Thomas, Felipe Lagos, Jose Borges, Agata Pacho, Emily Nicholls, Tom Turnbull and Joe Shaw. Special thanks go to Katherine Blanks who reminded me that dyslexia is both a curse and a strength and last but not least, Lilith Whittles who pitched in with some technical fixes and very necessary proofreading; her support and the support of other wonderful friends is incalculable.

This thesis is dedicated to my parents Michael and Lillian.

ABSTRACT

This thesis considers the role of online platforms (Wikipedia, Facebook, Twitter, etc.) in digital social research from a Science and Technology Studies (STS) perspective and proposes new conceptual, methodological and visual tactics, drawing on a series of empirical case studies concerning controversies over nuclear power.

Recent work in STS seeks to map science controversies (GM foods, nanotechnology, climate change, etc. Venturini 2010) using digital tools, which repurpose online platforms for social research (Rogers 2009). Yet these platforms not only provide data about controversies, they may also intervene in them as well. I propose that this requires studying them 'in action', drawing on the techniques of controversy analysis (Latour 1987) and actor-network theory (ANT). However, this research presents several challenges. How to delineate a study when controversies transcend particular platforms? How to define what is relevant when these platforms have their own relevance-defining metrics? How to track information flows within or between platforms?

The central argument of this thesis is that while researchers should capitalise on the affordances of these platforms, they must diverge from them as well. Theoretically, this means maintaining a tension between studying controversies and studying the platforms themselves. Methodologically this means decoupling methods from platform data structures: scraping less obvious data, juxtaposing quantitative and qualitative traces and presenting data in novel ways. Over three case studies, I will develop a series of mapping techniques for analysing controversies which *qualify the quantitative and make the less calculable more calculable,* revealing imbalances in the articulation and dissemination of controversies online which would remain hidden to platform-specific or qualitative approaches on their own. These exploratory techniques, which draw on work in the sociology of scientific representations (Woolgar and Lynch 1992), have implications for debates about big data, digital sociology, media studies and the relationship between quantitative and qualitative methods.

LIST OF FIGURES

Please note that detailed technical information and settings for each visualisation are provided in the Technical Appendix at the end. Since all of the original visualisations are presented as exploratory, working documents, rather than finished pieces for communicating findings, they have been made available, via a shared Dropbox folder, for consultation as zoomable and searchable PDFs. These can be viewed through the Dropbox interface but for best results, download and use Acrobat Reader or open with Google Chrome or Firefox (which have PDF plugins). In Acrobat, zoom controls are at the top, in Chrome/Firefox, hover mouse over bottom right corner for zoom controls or 'pinch' on a tablet or phone. Use Ctrl-F to search for particular text. Please note that some files are very large and make take some time to load. Files are available in this folder: https://goo.gl/8J8Vdq, which should have been shared with you previously, and shortened links can be found in the text.

FIGURES

IMAGES

TABLES

TABLE OF CONTENTS

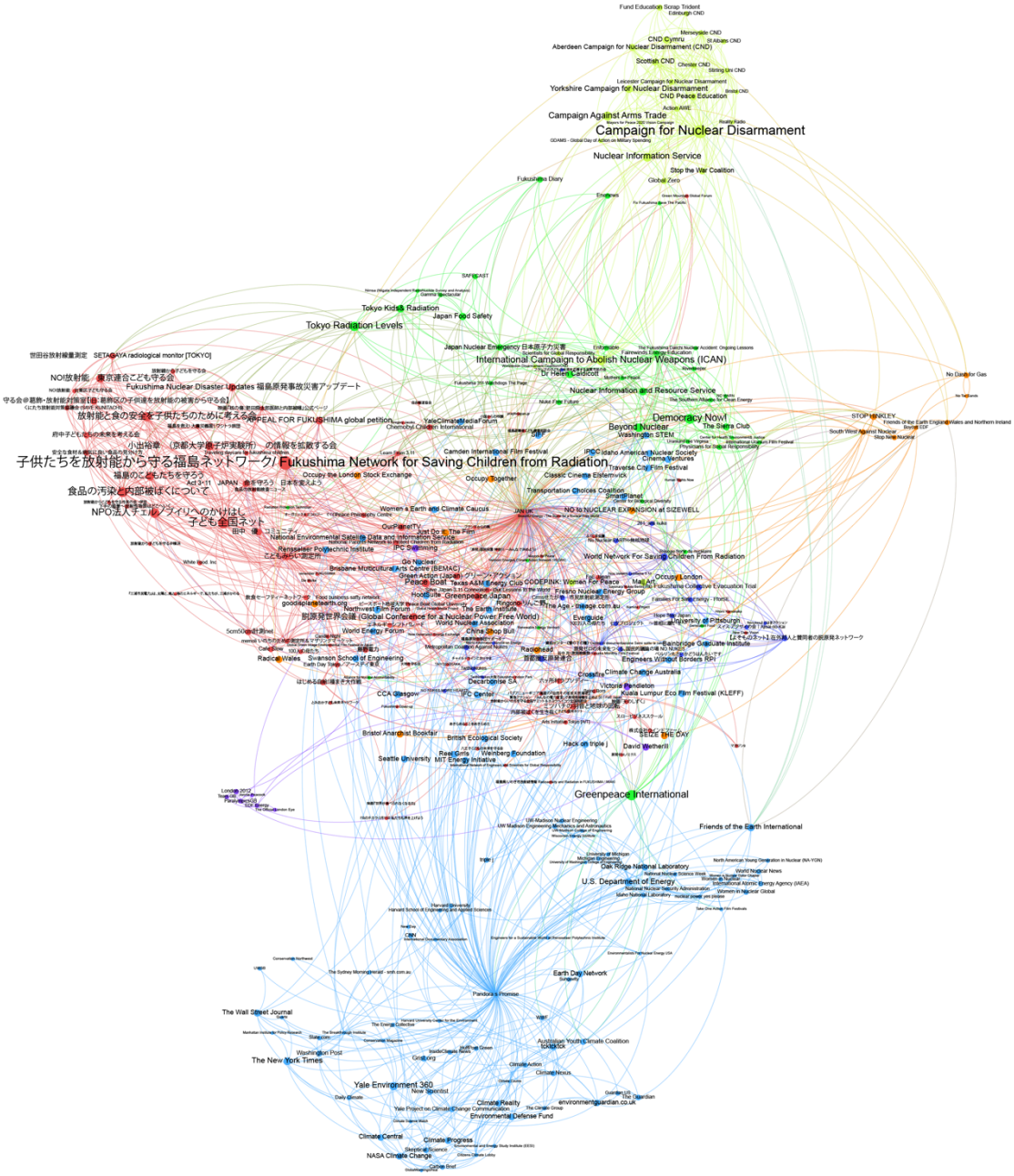


Figure 1. Like Network https://goo.gl/kZ8AWj: graph made using Netvizz (Rieder 2013) of Facebook pages related to nuclear power in the UK. The colours were produced with the modularity clustering algorithm, which attempts to group nodes based on their interconnectedness. Edges are curved clockwise from source to target.

I. INTRODUCTION

In an age of so-called 'big data' (Kitchin, 2014), automated tools and visualisations are increasingly ubiquitous. These techniques, which are increasingly made possible through the proliferation of Internet data, particularly through new online platforms and social media (van Dijck, 2013), have profound implications for social research. But what do these approaches really tell us?

The imposing tangle of lines and dots on the previous page (Figure 1) is a network diagram, produced with freely available data from the online platform Facebook. Facebook, which in 2014 boasted over 1.23 billion accounts[1], allows users to create personal profiles which are connected through mutually agreed friendship. It also allows users to create 'pages' for rock bands, corporations and activist causes which can be 'liked' by other users and pages in order to demonstrate their interest and follow updates. This user-generated data can be visualised as a 'like network'.

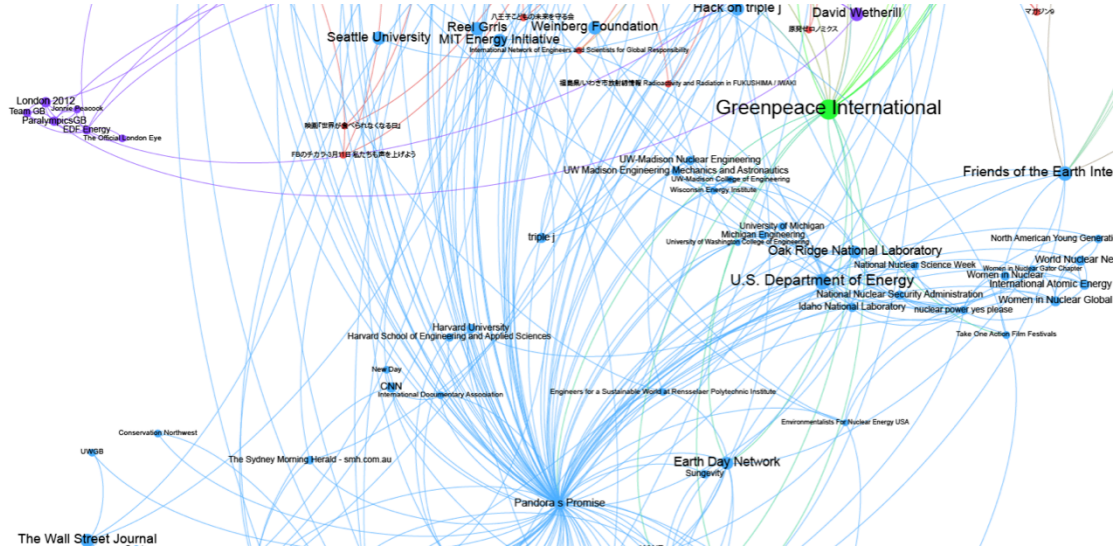
I will describe in Chapter V more precisely the rationale behind this visualisation but, briefly, I started with a researcher-defined selection of 'pages' related to nuclear power debates in the UK (in December 2013) and, using a digital tool called Netvizz (Rieder, 2013), obtained a list of all the pages 'liked' by that page and which of those pages in turn 'like' each other.[2] This information was pooled and visualised as a network of pages (each represented as a dot or 'node') connected by liking (represented as a line or 'edge' between nodes) using the popular network analysis tool Gephi (Bastian et al., 2009). I then arranged the nodes using a gravity-based algorithm 'Force Atlas 2' (Jacomy et al., 2014), so that nodes with more mutual connections are drawn together into clusters, which have also been automatically highlighted in different colours for ease of reading.

This map of Facebook pages appears to paint a picture of the current controversy over nuclear power plants in the UK, one which is to some extent corroborated through the

[1] According to *The Guardian* http://www.theguardian.com/news/datablog/2014/feb/04/facebook-in-numbers-statistics (Accessed 6 August 2015)

[2] I started with a list of pages related to nuclear power disputes in the UK (Boycott EDF, CNDUK, JanUK, South West Against Nuclear, Stop Hinkley, EDF Energy, Pandoras Promise).

literature, other online platforms, interviews and fieldwork.[3] Firstly, the network is dominated by pages focused on the 2011 Fukushima nuclear disaster. The cluster of mostly Japanese pages in red, which contains the densest inter-liking activity, speaks to the centrality of this event in galvanizing the anti-nuclear movement. It also shows how these pages engage with an established backbone of anti-nuclear opposition – mostly local groups centred around planned or existing nuclear plants – in orange, anti-nuclear weapons groups like Campaign for Nuclear Disarmament (CND), in yellow, and environmental groups such as Greenpeace and Friends of the Earth (FoE) in green.



Detail of Figure 1. Like Network, focusing on nuclear operator EDF and Pro-nuclear environmental film Pandora's Promise.

Predictably, French company EDF Energy, the sole nuclear power plant operator in the UK, has few links to or from the network: they are in the small purple cluster consisting of London 2012 Olympic pages and the London Eye, of which they were corporate sponsors.[4] The map also shows a possible tension within the environmental community over nuclear power, which is argued by some to be a 'low carbon' alternative to fossil fuels: a solution to climate change. This is expressed in the blue cluster in which the page for pro-nuclear environmental film *Pandora's Promise* 'likes'

[3] I hesitate to say corroborated because as this thesis will make clear, our methods produce very partial versions of a particular object such as this controversy, but while they are in dialogue they do not so easily converge on a reality independent of them.

[4] They are also linked to some of the more tangential pages in the network, which include various media sources and think tanks. They are only liked back by Victoria Pendleton and David Wetherill, two athletes in the 2012 Olympics of which EDF is also a corporate sponsor.

several environmental groups such as Friends of the Earth and Greenpeace, who do not return the compliment.[5]

Platform-based tools like the above are thought to greatly extend our ability to research complex phenomena like nuclear power controversies – from a short list of nuclear related groups and a few clicks of a button I was able to obtain a map of over three hundred groups and their relationships. Some have even argued that these new sources of data may help rethink divisions between quantitative and qualitative methods; macro and micro levels of analysis (Venturini and Latour, 2010), allowing for fluid zooming between individual cases and larger relationships without aggregation or reduction. When used critically, these tools also pose questions about the platforms through which they are generated. What does it mean for a page to 'like' another page? Why are some groups, such as Japanese Against Nuclear UK (JANUK) more promiscuous in their 'liking'? Is this equivalent to a referral or citation (Marres and Rogers, 2000) or are the page administrators using 'liking' to monitor the activities of other pages, including those of their opponents? Is 'liking' part of a strategic and self-conscious game of boosting visibility and rankings in a saturated social media landscape? While on one hand this is a map of a nuclear controversy it can also be read as a map of divergent tactics of *publicizing* the controversy.[6]

However, there is also much that the above map conceals, or only alludes to, because it is inextricably bound to a *particular* online platform and its data structures. Several of the pages represent other media outlets: a liberal public television program, anti-nuclear blogs, mainstream newspapers; how do these more traditional media relate to activities on Facebook? Is there some unseen bridge between the pro and anti camps on other media channels? What about the numerous discursive statements, images and other practices on these platforms, which, qualitative researchers may protest, cannot be so easily summed up by a 'like'?

This thesis is about addressing the role of online platforms such as Facebook in online social research in both conceptual and methodological terms. These platforms make available large amounts of data and enable push-button digital tools like the above but they also may shape research aims and possibilities. However, this is not another

[5] Only a couple of pages in the network 'like' the film back: a documentary film association and The Breakthrough Institute – an environmental think tank known for being pro-nuclear.
[6] Or is it a map of my research design? Crawler type tools may be heavily dependent on their starting points and might change dramatically with different settings.

attempt to dismiss digital tools or platform research from the perspective of qualitative or ethnographic research. Neither is it a treatise on design or an elaboration of existing quantitative approaches – I have neither the technical capabilities nor the design sense to make such a contribution. It instead advances an approach to data in which researchers should capitalise on but also diverge from the data formats and methods implied by platforms. This requires different sorts of tactics and visualisations and a rethinking of the domains of quantitative and qualitative work.

This work emerges from a very specific set of problems, situated at the thriving intersection of Science and Technology Studies (STS) and media studies: two disciplines that are currently the subject of important new collaborations (Gillespie et al., 2014). In particular, this thesis contributes to work that deploys online media (hyperlinks, search engines, Wikipedia, Facebook, Twitter, Instagram) and digital tools to map public science controversies such as those over GM foods, climate change, nanotechnology and nuclear power (Venturini, 2010b). Controversies have long been a focus for STS, but tracking these complex, roaming and volatile objects has increasingly prompted researchers to turn to the web (Marres and Rogers, 2000, 2005) and more recently online platforms: new websites, sometimes associated with Web 2.0 (O'Reilly, 2005) or social media (van Dijck, 2013), many of which are driven by user contributions.

However, these platforms do not just provide data about controversies, allowing researchers to identify key players and the terms of the debate, they also *mediate* access to them and, according to recent literature, may intervene in them as well. The key players will actively be using these platforms to promote themselves and their causes and cast their opponents in an unfavourable light, which has consequences for how debates play out. In other words, these platforms are neither neutral nor separate from the controversy.

What this requires, I suggest, is that researchers should not only use online platforms to follow controversies but also study these platforms 'in action' – as *venues* through which the controversies play out (Latour, 1987). When I say 'in action', I am invoking a more qualitative, ethnographic tradition in STS, including the technique of Actor-Network Theory (ANT) and the 'engaged' tradition in Public Understanding of Science (PUS) which have examined science controversies through the analysis of material settings, such as the laboratory, texts, such as scientific journals, and later through

public hearings, protests and policy documents. This work revealed some of the politics and asymmetries of resources underlying these seemingly technical matters and modes of engagement. These techniques could similarly be used to demonstrate how online platforms, as arenas in which competing groups jostle for position, may benefit certain positions at the expense of others, in ways that shape the controversy *and* digital maps of it.

However, importing these mainly qualitative, small-scale techniques online is no simple task. The size and complexity of these new data sources may require qualitative researchers to embrace some form of quantitative analysis or mapping. Researchers can helpfully draw on techniques known as 'digital methods' (Rogers, 2009, 2013b) such as the 'like' network above which repurpose these online platforms as research tools and highlight some of the 'politics of platforms' (Gillespie, 2010). Yet several problems remain for both quantitative and qualitative approaches:

- How to delineate a study when controversies transcend particular online platforms?
- How to define what is relevant for the study when these platforms have their own relevance-defining metrics, which not only measure, but shape interactions?
- How to track dynamic information flows within or between platforms?

The central argument is that to study controversies 'in action' we need to capitalise on the affordances[7] of these platforms, which I will later refer to as 'devices' (Law and Ruppert, 2013) but also resist their pull. Theoretically, this requires de-centring the object of study, maintaining a tension between studying platforms and controversies which transcend particular platforms. Methodologically, this means not only analysing the readily quantifiable traces offered up by platforms, but also discursive texts, unstructured data and activities which fall outside their nominal boundaries, recombining this data in unconventional ways.

Through a series of empirical case studies concerning nuclear power debates on various platforms (including Wikipedia, Facebook and Twitter) I will develop a set of

[7] I deliberately use 'affordance' here to mean not just what they enable, but acknowledging, following Mike Michael, different ways they may interfere with or mediate relationships (Michael, 2000) in this case between methods and objects like controversies.

techniques and data visualisations which work in tandem with qualitative, textual analysis inspired by STS. They will *qualify the quantitative and calculate the less calculable.* In doing so, they will reveal some of the politics of how these platforms represent and possibly impact controversies, which may be missed by traditional qualitative techniques or platform-specific tools on their own. The hope is that these techniques that emerge from the unique object of controversies and STS approaches to (social) scientific representations will also be relevant to wider debates about social media research, digital sociology and the value of big data techniques in the social sciences more generally.

But before getting into these theoretical and methodological debates about the role of platforms in digital research, I want to step back and explain why the intersection of public science controversies and online media is an important topic and why it warrants a new approach. I will first introduce public science controversies as an object and using the case of nuclear power, explain what is at stake in their analysis. In the remainder of this chapter I will propose that these new online platforms not only represent a new source of data for observing controversies, they represent a viable object for controversy analysis in their own right and this requires more sustained analysis and deeper engagement between controversy analysis, media studies and new digital tools.

However, I suggest that there is an ambiguity in how the role of online platforms is conceptualized. Are they sites of knowledge production? Are they technologies for assembling oppositions? Are they forms of alternative media? Although STS has an established language for talking about the former two, it has an ambiguous relationship to media and this leads to several conceptual and methodological issues which will be addressed over the course of this thesis.

01. PUBLIC SCIENCE CONTROVERSIES

Why Controversies?

A central concern of social scientists and political theorists for the last 30 years has been the role of technical decisions in modern democracies (Winner 1989). According to Ulrich Beck (1992), politics as we know it has been displaced from traditional

political arenas, such as parliaments, to back-room negotiations between experts and scientists. Beck calls this 'sub-politics' – a version of politics as problem solving, which often takes place within institutions away from public scrutiny.[8] Beck theorized that modern science was manufacturing 'risks', unforeseen and difficult to define consequences of human activity such as pollution and radiation. In their attempts to manage and make sense of these risks, scientists and experts (as well as the media, Beck, 2000) claiming to detect and make risks visible are in privileged positions in the 'risk society'. For political theorist Stephen Turner (2003), the position of 'experts' reveals contradictions in conceptions of liberal democracy, because it means that certain individuals have more voice than others (see also Hilgartner, 2000).

Occasionally these ostensibly technical matters burst into view as public controversies, such as those over GM foods (Jasanoff, 2005), nanotechnology (McCarthy and Kelty, 2010), global warming (Latour, 2004), chemical spills (Jasanoff, 1994) and air pollution (Barry, 2001). These matters are difficult to settle or comprehend because they straddle the fictive walls our 'modern' society has erected between science and politics (Latour, 1993). In public, controversies enable new modes of doing politics (Marres, 2005b) but are just as susceptible to capture by de-politicizing procedures (Barry, 2005). Over the last 30 years, these controversies have become central empirical materials for the field of STS and PUS, particularly strains inspired by ANT.

Public science controversies are an important object of research because they represent a limit-case for modern political institutions (and contemporary political theory). Yet they are also important because of certain analytical and methodological capacities, which have implications for the social sciences more generally. According to researchers associated with ANT, controversies offer occasions where the social breaks down, exposing its component parts and processes. Scientific knowledge and technical processes are often 'black boxed': their conditions of production are occluded from view (Latour, 1987). But in controversies these more or less fragile alliances, which

[8] Similar work refers to this phenomenon provocatively as 'post politics' (Johnstone, 2014) which implies the active suppressing of 'dissensus' or disagreement. In some ways, this assertion is based assumptions about what proper politics should be (parliamentary representative or deliberative democracy). Marres notes that this misses many of the less obvious modalities in which politics happens (2012a). Barry, helpfully defines a sliding scale in terms of 'politics', which he defines as a space of disagreement, and 'anti-politics', the suppression of disagreement, but contends that politics and anti-politics can occur in both traditional locations of politics (parliaments, hearings) or non-traditional spaces (Barry, 2005).

underwrite the legitimacy of knowledge, technology and institutions are made available for analysis.

Controversies also enable researchers to short-circuit dualisms such as those between realism and social constructivism. Proponents of broadly constructivist approaches have argued that while Beck correctly diagnoses a key problem in contemporary society, he does not go far enough because he retains a 'realist view of science' (Lash et al., 1996; Wynne, 1996, 2005). He sometimes speaks about 'real risks' as opposed to public perceptions of risks, while at other times the risks and their constructions within science are inseparable.[9] Wynne and others advocate not deciding *a priori* which participants in a controversy (such as experts or laypeople) have privileged access to nature or reality. But advocates of ANT are also critical of 'constructivism' which often uses social structure or culture as explanatory tools, when in ANT the stability of social institutions and knowledge is what must be explained. Controversially, this not only involves social actors – non-human technology and nature are also implicated in the simultaneous construction of knowledge and society. So, while all knowledge is constructed, some knowledge is more 'real' in that it can survive multiple trials (Latour, 1999) – it is *better* constructed. Controversies are privileged objects in that they make these trials visible.

So one way to think of the approach of this thesis is to not only take on controversies as a *topic* to be apprehended through online media and digital methods but to also follow through on controversy as an analytic and methodological *trope* in which everything, from the boundaries of the study, the actors involved and the methods employed, is at stake in the analysis.

Why Nuclear Power?

To explore these methodological concerns with online media and public science controversies I could have chosen any number of controversies to study.[10] Nuclear power presented itself as an option when, in 2013, on the second anniversary of the

[9] Beck later acknowledged that he saw realism or constructivism as a pragmatic choice to be taken up in the service of his argument: for example constructivism to problematise scientific representations, realism to make prescriptive claims (2000). However, ultimately in his account it is a reconfigured science which must be the solution to the problems it generates, and this effectively reifies the divide between scientific (expert) and lay forms of knowledge.
[10] In fact, nuclear power arrived relatively late in the proceedings.

Fukushima nuclear disaster, I found images of widespread direct action splashed across my social media feeds. The protesters linked the recent catastrophe in Japan to the restarting of Britain's nuclear programme, while the nuclear industry presented the programme as an urgent response to climate change (Peoples, 2014). Just two weeks later, planning permission was granted for the first nuclear power plant in Britain in a generation.

Aside from being topical, there are two reasons why nuclear is an appropriate case for this project: it is historically a 'hot' controversy, and thus easy to analyse, and it has been at the forefront of battles over public participation, something that online platforms are supposed to redress. It also raises some questions about the role of various media in controversies, the significance of which will become clear later.

According to Brian Wynne, nuclear is '…the most iconically controversial of modern technologies' (Wynne, 2011: 1). In the language of ANT, controversies are more easy to study when they are 'hot', when the social is molten and malleable as opposed to when they are 'cold' and increasingly solid (Callon et al., 2001; Venturini, 2010b).[11] Dorothy Nelkin (1981) cites several reasons for this: their scale[12] – that is, the scope of possible economic and environmental impacts and the involvement of powerful actors and institutions (governments and industry); their complexity – the seeming inevitability that the complex technology will result in failures, contaminations and accidents (Perrow, 1984); and finally, their unpredictability – or the lack of consensus over measurement systems for determining risks or effects in a clear way.[13]

The second reason why nuclear was an apt controversy to study was that nuclear has been central to the development of accounts of public participation in science. STS and PUS scholars have in various ways argued that 'the public' or laypeople, who possess other valuable forms of knowledge, must be allowed more participation in

[11] As I will discuss later, while the Fukushima disaster itself was unequivocally 'hot', nuclear debates in the UK have been merely simmering. What then becomes interesting about, particularly 'social media', is their seeming ability to make things 'hot'.

[12] In today's terminology it would be preferable to think of scale not as a given – as in 'the national scale' – but rather as an achievement of the extending of networks across locales (Latour, 2005).

[13] In line with Beck's argument, ionizing radiation, the kind produced by nuclear power plants, is particularly resistant to scientific attempts to domesticate it. It is hard to study in a laboratory because it is so dependent on local environmental conditions like wind and soil absorption (Wynne 1992). Also, while the effects of external absorption are mostly known, the effects of internal absorption, through soil, animals and food are an open debate within science (Dorfman et al., 2012).

controversies over technical decisions. But according to this literature past attempts at participation have been largely confined to experts, and attempts to involve the public in the proceedings have been mostly tokenistic (Nelkin, 1971, 1974). When governments have granted public consultations over the siting of plants, the terms on which participation happens: who counts as an expert or legitimate participant and what kinds of evidence are admissible have been unnecessarily restrictive (Welsh 2003). According to Brian Wynne (2011), who studied the Windscale Inquiry over the development of the THORP nuclear reprocessing plant, because the controversy was phrased in terms of risk, no other relational, affective or social premises were allowed. Wynn calls this 'legalism': the assumption that 'issues can be purified into precise empirical questions' and that political disputes could be solved with propositional claims.

But nuclear also provided these critics with one of the clearest cases for increased participation. In Wynne's account of Cumbrian sheep farmers seemingly affected by Chernobyl fallout (Wynne, 1992), the scientists on site had to revise their predictions about long-term risks because they ignored the farmers protestations about the rockier than average soil. The universalising impulse of science often misses the local particularities as well as other ways of knowing.[14] Despite increased gestures and more open settings of participation, it is still the case today that many key decisions about nuclear, in the UK at least, happen in policy circles outside of the public view (Aubrey 1991, Johnstonne 2014).

So, nuclear power controversies have long provided rich empirical material to researchers in STS and related disciplines because, according to this literature, nuclear controversies are often accompanied by democratic deficits. Because of the extremely complex science and engineering involved and the types of powerful actors enrolled, experts and technocrats have dominated the discussions at the expense of concerned lay actors. STS is unique in its presentation of this problem because the content of scientific knowledge is analysed together with the politics of different modes of assembly and participation. As Latour phrases it, STS is concerned with the relationship between 'representations', presentations of facts or knowledge, scientific or otherwise and 'representation' in the political sense: forms of politics, of speaking

[14] As Wynne puts it 'This kind of analysis of scientific knowledge as constructed and used in risk and environmental issues opens it up to recognisable indeterminacies as to whether the controlled and artificial conditions assumed in the analytical process (perhaps the laboratory) will actually prevail in practice, so it is implied, everywhere at all times.' (Wynne, 1996: 58)

on behalf of others (Latour and Weibel, 2005). The reason I emphasize this is that while these two facets have always been analysed together and were present in the early interventions by STS scholars in the internet (Marres, 2006), something about the rise of platforms seems to drive a wedge between these two modes of analysis. In the next section I will propose that new online platforms may offer new venues or technologies through which lay actors can both intervene in knowledge and politics, and this requires more sustained analysis.

Now, there is another key player in these controversies, which is dealt with less explicitly in STS accounts above and that is: the media. When I refer to 'the media' as opposed to 'media', I generally mean mainstream media, specifically journalism, in common usage, though in the next chapter I will specifically question this understanding.[15] In some cases the media makes visible public opposition through opinion polls and provocative cover stories (Nelkin, 1974), in other cases it becomes a flashpoint for disputes about the proper domain of science (Nowotny and Hirsch, 1980). The media has been instrumental in publicizing accidents such as Three Mile Island and Chernobyl (Friedman, 2011; Otway et al., 1988) which would otherwise be concealed by the industry and governments (Hilgartner et al., 1983) and this helped to mobilise local groups and connect them to related struggles around the world (Welsh, 2000). The media may also shift the terms of the debate within participatory settings (Aubrey, 1991), or contest expert representations of accidents (Luke, 1987), even when they are not in a position to contest the science itself. Yet, while the media has played an important if ambivalent role in controversies, it also historically represent another arena in which lay actors can only participate in circumscribed ways (such as opinion polls and letters to the editor). The reason I bring this up is that, as I will discuss in the next chapter, STS has not always addressed the role of the media, or communications technologies more generally, in controversies in a systematic way, though new bridges are currently being formed in the literature. As I will suggest in the next section, this deeper engagement between scholars of controversy and media research may also be necessitated by the emergence of new forms of online media.

02. THE PROMISE OF ONLINE PLATFORMS

As nuclear controversies have largely disappeared from the front-pages and the official

[15] I will also in the empirical chapters attempt to specify particular media outlets rather than refer to the media in this way as a monolithic whole.

participatory process, in the UK at least, has been dismantled, other avenues of participation have multiplied on the internet. In this section, I want to raise the possibility that new online platforms may offer new settings or avenues for so-called 'lay actors' to intervene in public science controversies, first through the emerging literature and then through the specific example of the Fukushima disaster.

What are platforms?

We are often told, breathlessly, of the emancipatory potential of new technologies on the web sometimes associated with Web 2.0 (O'Reilly, 2005), new media (Livingstone and Lievrouw, 2009) or more recently 'social media' (van Dijck, 2013). Beer (2009) describes these technologies which include social networking sites (Twitter, Facebook), Wikis (Wikipedia), blogs and Folksonomies (Youtube, Flickr) examples of 'participatory web cultures' which are driven by user contributions. These platforms are often based on models of collaboration and 'openness' or transparency emerging from Free, Libre and Open Source Software (FLOSS) culture (Tkacz, 2014).

Although there is only space to glance over them, there have been a flurry of studies from outside STS, mostly in media and communications but also law, policy and cultural studies which have attempted to theorise the significance of these new technologies and practices. Significantly, many of these authors phrase the novelty of these media or technologies in relation to traditional broadcast media as do certain accounts of Fukushima (Friedman, 2011). Benkler (2006), for example contrasts the intensely hierarchical 'few-to-many' logic of broadcast media to the distributed multi-centric logic of networked media. It is not just pre-ordained experts and professional journalists who have a voice, now bloggers, citizen journalists, citizen scientists and everyday users who can produce content (Shirky, 2009; Tapscott, 2006). These media are thought to enable greater participation in protests and social movements (for example Castells, 2013; Mercea, 2013; Thorson, 2014). For some, these new modes of participation herald a revitalised public sphere (Papacharissi, 2002) allowing bloggers, social media users and commenters to respond to dominant voices as 'produsers' (Bruns, 2009) or citizen journalists (Allan 2006). Axel Bruns' book *Gatewatching* (2005) argues that platforms like blogs are increasingly involved in selecting and disseminating sources, in a reversal of the classic editor as 'gatekeeper' (White, 1950) understanding of media. These media are also associated with new modes of circulation such as 'sharing' (Van Dijck 2009) and the 'viral' (Kullenberg and Palmaas,

2009) spread of information.

STS informed researchers, however, have been somewhat slower to attend to these new phenomena, possibly because in relation to the meaty materials of science, economics and engineering, new media may seem hopelessly banal or slight (Rogers 2013). There have certainly been STS studies of the internet, which often view it in a broader historical perspective, such as Fred Turner on the emergence of 'cyberculture' out of networks of countercultural figures (Turner, 2010), Bowker and Star on informational infrastructure (Bowker and Star, 1999) and a famous edited volume confronting internet hype (Woolgar, 2002). But it is only in the last couple of years that these online platforms are receiving sustained attention from STS inspired researchers (Gillespie, 2010; Langlois et al., 2009; Lievrouw, 2011; Rogers, 2013a; van Dijck, 2013).

As Boczkowski and Lievrouw remark in their contribution to the *STS Handbook* (2008), writers from STS are naturally reluctant to make definitive claims about causality or the impact of new technology, wanting to avoid technological determinism and social constructivism. In various ways the above authors all advocate interrogating how research subjects empirically define these phenomena rather than imposing their own theorisations. According to Gillespie (2010), who studied some of the popular literature and business manifestos of new tech companies, these new technologies are collectively referred to, not as websites, but as 'platforms': rhetorically suggesting both the *elevation* of everyday users and that these users are placed on a *level* (non-hierarchical) playing field. But as Gillespie points out, these definitions gloss over much of the politics and imbalances, which favour certain sorts of content and participants over others.

I will refer to these technologies, which are broadly interactive and larger and more complex than most discreet websites and blogs from now on as 'platforms', keeping in mind that it is precisely the attendant claims of enhanced participation and neutrality, which need to be evaluated.[16] I prefer this rather blank term to 'social media' or 'new media' or 'Web 2.0' because, as I will show, it is important to not decide in advance which websites or technologies are most consequential or which are most

[16] I am using this term as a blank infralanguage to distinguish the object of study from actual invocations of terms like 'new media' and 'social media' in the world (see Slater, 2014) which act on the world and may themselves become part of the study. Some of the above authors investigate the use of these terms specifically but what is more relevant for the current study is the more implicit understandings users have of what these platforms are for.

'participatory'. When I use the above terms I mean them in the industry definition. However, because I am looking at these platforms through a handful of case studies rather than from a broad historical vantage, I will not be making any definitive claims about their capacities for politics or participation but rather develop approaches through which such evaluations could be made in the future.

Although the general literature points to possible shifts in the role of the 'expert', increased participation and the wider circulation of information that should be relevant to public science controversies, the potential of platforms becomes much clearer through the lens of a particular controversy.

Fukushima and Platforms

In March 2011, an earthquake of magnitude 9.0 and the resulting tsunami knocked out power to the Fukushima-Daichii Plant in Japan, which triggered a series of partial meltdowns, explosions and radiation releases over the coming days, causing the government to evacuate residents in a 20-mile radius. Initially, in Japan, this event caused a large public outcry and the shutting down of all nuclear plants (some have since been turned on). In Germany the event accelerated the scaling back of nuclear in favour of renewables. Although the incident was marked by asymmetries of resources and information much like previous nuclear accidents (Three Mile Island, Chernobyl); Fukushima played out quite differently, at least to some extent due to the presence of online media (Friedman, 2011). Although much has been written about Fukushima already, I will confine the discussion to two STS informed studies that directly connect science controversies and online platforms, which make two key interrelated claims.

As Morita, Blok and Kimura (2013) recount, in the first few hours, the government (and mainstream media relying on them) downplayed the potential severity of the incident. But an explosion at Reactor 1, the following day seemed to contradict this message. In the days that followed, distrust in the government grew as they mishandled the SPEEDI radiation monitoring data – several key days of readings were mysteriously lost.[17] When radiation readings were supplied by the plant owner TEPCO (Tokyo Electric Power Company), they were often in scanned PDFs, not machine readable for easy analysis or viewing on smart phones. Distrust of the official account was also fuelled by

[17] This was possibly due to the same flooding and blackouts which caused the incident in the first place.

discrepancies between domestic and international reporting of the disaster.[18]

In response to this growing uncertainty, a lone Twitter user (@MFkurochan) created a layer on the open platform Google Maps where live radiation readings could be uploaded, in theory by anyone with an internet connection. This map grew in complexity over the coming days and weeks as both amateurs and scientists published their readings. As discussed earlier, normally the authority of scientific representations is enforced through the 'black boxing' of the conditions of production (Latour 1986) but in this case, as Morita, Blok and Kimura argue, revealing the radiation infrastructure (pointing out the location of monitoring posts and even taking pictures of them) lent *credibility* to the representations, regardless of the credentials of the person uploading the readings. This case gives a very clear example of what STS researchers have called 'research in the wild' (Callon et al., 2001) that is knowledge generated in controversies by non-scientists using different, but not illegitimate, methods. Online platforms like Google Maps and USTREAM thus facilitated new forms of knowledge production when existing institutions and experts could not.

Another STS-informed study (2011) describes how several citizens collaborated in sourcing and cleaning the available government radiation data, mapping the results and creating 'data mashups', in order to construct a more cohesive picture of the accident. Jean-Christophe Plantin argues that these maps helped assemble potentially affected actors, who would have otherwise been isolated and fragmented. Plantin is drawing on Marre's account, discussed in the next chapter, of how objects and technologies can organize 'publics' (2005a) or particular modes of participation (Marres, 2012a). So just like public hearings, opinion polls and other 'technologies of elicitation' (Lezaun and Soneryd, 2007) it seems that online platforms have the ability to forge novel, ad hoc modes of assembly in the absence of institutional technologies and settings.

Both of the above claims are about very particular platforms, both associated with mapping practices. However, I also want to draw attention to a wider claim about Fukushima and online platforms which is present in the above accounts and frequently made elsewhere: that the presence of participatory media gave lay-actors means of contesting the nuclear industry and government articulations of the disaster (Slater et

[18] Also NISA and TEPCO press conferences were held separately and releases often contradicted each other.

al., 2012).[19] Morita, Blok and Kimura also describe how 'Widely divergent stories about the disaster started circulating on the Japanese-language internet, exerting a persuasive force on people…' (Morita et al., 2013: 84). The effect of these interventions was that a significant share of media representations (in particular, those outside of Japan) were able to present the incident as a systematic failure of the nuclear industry (Yamamura, 2012) and the government, rather than as 'an act of God' or the result of individual human error: the preferred explanations of nuclear accidents in the past (Luke, 1987). Now, STS scholars may have good reason to shy away from such broad claims about online media influencing old media in this way, but it is still an important potential of these new technologies that should be entertained.

In this section I have tried to offer the possibility that online media can intervene precisely in the ways that have been lacking in the past: they present new modes of knowledge creation and at the same time new modes of assembly and representation both of which exist outside of institutional venues. So while they certainly offer new sources of data for visualizing and mapping controversies, they also present new venues, technologies and practices through which controversies may play out. However, this all depends on the terms on which we understand their potential contribution.

03. RESEARCH CHALLENGES

Even if only some of the more utopian claims made about platforms in the wider literature are true, then these platforms could provide new avenues for laypeople or publics to participate in public science controversies, as they did in the Fukushima disaster. Yet researching controversies on participatory media presents several daunting challenges: methodological but also conceptual.

As Christopher Kelty points out, there is a profound ambiguity in the literature about what is meant by 'participation' in relation to these platforms (Kelty, 2013). Kelty has in his crosshairs a much cited Malcolm Gladwell editorial about the fate of activism online (2010). Gladwell argues that social media is more horizontal and democratic,

[19] This is not to imply that there was an information free-for-all, but the presence of new participatory platforms meant that institutionally recognized experts and journalists were not the sole arbiters of this unfolding reality (Slater et al., 2012).

but affecting real social change, requires organisations with rules and hierarchies (such as those central to the American civil rights movement).

> 'Social networks are effective at increasing *participation*—by lessening the level of motivation that participation requires.' (Gladwell, 2010)

But in the case of online platforms is this participation in social movements? Or participation in the state and government? Or participation in media discourse? Or the latter contributing to the former? According to Kelty, it seems as if these possibilities have become conflated without understanding the mechanisms by which, for example, a media intervention affects changes in the state. In the much discussed case of Egypt, during the 2011 Arab Spring, for example, were the revolutionaries in Tahrir Square using social media like Facebook and Twitter to organize their members (Gerbaudo, 2012) in direct action or were they, through millions of 'likes' and Tweets making themselves and their messages visible to the mass media, or contesting mainstream media representations (Meraz and Papacharissi, 2013). Which of these tactics resulted in the Government effectively switching off the internet and President Mubarak resigning?

The same thing could be asked again of the Fukushima example: were online platforms producing new knowledge, like a kind of alternative laboratory, or were they creating new spaces for participation, or providing alternative articulations of the controversy vis-à-vis the mass media? Which of these shifted the narrative and caused the backlash against the nuclear industry? Answering these questions is difficult and surely the answer is to some extent all three, but each of these possibilities involves approaching online platforms with different conceptual equipment.

Researchers could trace the construction of facts on platforms through networks of actors and technologies, asking how particular forms of knowledge or expertise become legitimate. One could also study platforms as if they were a participatory setting parallel to STS studies of public hearings. This would involve interrogating and evaluating the terms and limits placed on participation and how this is enforced through the specific affordances of platforms. However, the problem is that these platforms are not quite like technologies of elicitation nor laboratories. They circulate information in different ways, and it is by no means clear that 'facts' or knowledge are the primary thing that travels. In short, understanding the contribution of platforms

means attending to their specificity and this might require interrogating them in relation to media.[20]

However the choice to analyse participatory platforms as media or as technology or indeed as a tool for social movements is not just an analytic decision but an empirical *topic* which the participants wrestle with. Is Wikipedia an encyclopaedia or a news source? Is Facebook an organising tool or a sort of megaphone for social movements? Is Twitter a broadcast medium or an interactive forum conversation? I argue throughout this thesis that we need to keep the definition of platforms and their potential contribution open.

But while STS is well equipped to deal with participatory platforms as sites of knowledge production or participation, it does not have a strong framework for dealing with them as media. Now granted, from an STS perspective there is nothing categorically separate about media as a class of phenomena in relation to other forms of representations and technologies, but media involves distinct modes of mediating and circulating, which must be attended to in their specificity, not reduced to pale shades of science through comparisons with laboratories and scientific papers. However, unlike the flow of information through scientific texts, the wide reach of broadcast media makes it difficult to trace, and analyses are often fragmented between producers, texts and audiences.[21]

Yet online platforms may finally make the STS analysis of media effects tractable, because flows of information have become to some extent visible, folded into these platforms which are extensively archived and woven through with metrics, with algorithms silently selecting and ranking content from vast databases concealed in the back end. However, the fact that these platforms are so well documented and (potentially) analyzable represents an opportunity but also poses several challenges.

This is partly the case because of inherited divisions between quantitative and qualitative approaches. The traceability and structuredness of this data has prompted

[20] The involvement of media may also explain, for example, certain divergences between cases. Why are certain accidents downplayed while others are heavily scrutinized? Why does essentially the same technology become (mostly) a source of national pride in France (Hecht, 2009; Touraine, 1983) and a political lightning rod in Indonesia (Amir, 2009)?

[21] There are also classic media questions regarding the audience and reception of media messages and symbolic and affective readings of media texts, which ANT, and thus this thesis will have to remain silent on (Couldry, 2008).

the rise of quantitative, 'big data' techniques which purport to analyse social life at an unprecedented scale (Kitchin, 2014). This turn to data-driven approaches has led to a flurry of non-empirical theorizing and the closing ranks of traditional qualitative social research methods (interview, ethnography etc.), both of which have generated significant insights and critiques of the computational vanguard (e.g. boyd and Crawford, 2012).

STS approaches, in contrast have long embraced digital tools in tandem with qualitative work (Callon et al., 1986; Latour et al., 1992) even in relation to studying controversies on the internet (Marres and Rogers, 2000, 2005) but online platforms, which provide formatted data for research through Application Programming Interfaces (APIs), may exert even more influence on the study (boyd and Crawford, 2011; Marres and Weltevrede, 2013) than even web data. In particular, I want to raise three interrelated problems facing research on and with platforms which will be addressed in the empirical chapters.

1) One of the formative findings of ANT ethnographies of the laboratory is that studying controversies might mean following actors *outside* the laboratory. Controversies might, similarly, lead us outside of particular platforms revealing wider networks of interlinked actors and technologies. Yet, while the data on platforms, particularly hyperlinks, allows for this, platform data formats may complicate this tracing. Also, digital tools are mostly, by design, platform-specific. In any case, feeling out the indeterminate boundaries of the study is a difficult task.

2) Much of the formatted data available from platform APIs comes in the form of metrics or quantifiable data, which can be ranked and graphed. On one hand, it has been argued that these traces may bias the research toward the most popular or 'trending' content – the platform's definition of what is relevant – which may diverge from what is sociologically interesting (Marres and Weltevrede, 2013). But in addition, these rankings not only measure, but also shape the interactions under examination, as users attempt to reflexively game the rankings. How are we to empirically study the effects of this quantification without becoming beholden to it?

3) Finally, as alluded to above, one of the main barriers to an STS understanding of media is a working understanding of how information flows either between, or within, platforms. The reason why it is difficult to claim shifts in the 'narrative' of Fukushima is

that we cannot trace the dissemination of particular contents from producers to audiences; the effects are *presumed*. While this is a perennial problem of media studies, which I do not hope to resolve here, platforms may allow some limited insights into this process because of the amount of feedback and exchanges they make visible. The barrier is again our existing assumptions about information diffusion and the way platforms format data.

This thesis is about addressing these sorts of problems. What frame of reference to approach platforms with? How to delineate the study? How to proceed methodologically when our methods are built on the platforms themselves? These problems need to be addressed in order to more confidently proceed with the task of evaluating the promise of new participatory media in science controversies. Fortunately, these platforms provide all the necessary equipment to overcome these challenges, if only we can un-think these conceptual hang-ups and methodological baggage.

04. OUTLINE OF THE THESIS

In this introduction chapter I made the case for why it is interesting to study new online platforms in relation to public science controversies. While these platforms disclose data which may be useful for monitoring and mapping controversies, platforms, according to the literature, hold the promise of intervening in controversies as well.

I discussed the topic of nuclear power as a particularly extreme case of a public science controversy, particularly one whose outcomes seem to be wrapped up with the role of the media. The STS literature on nuclear tells a story of democratic deficits and the exclusion of affected people and laypersons from technical decisions. Although instances of participation have increased over time, they have often been circumscribed in terms of who can participate, and with what sorts of evidence, which are set in advance. I then brought this literature in to dialogue with literature on online platforms which suggests that these technologies might upset traditional imbalances between experts and laypeople, institutions and publics. Particularly in the case of Fukushima, these platforms enabled new voices to contest the official account given by the government and industry. I also cautioned that that evaluating the contribution of

these platforms to controversies depends very much on how participation is defined, and proposed that perhaps their most important interventions is in the media landscape. Yet as I explained, this requires that STS confronts the role of media in controversies.

In the next chapter (Chapter II), I will think through what would be entailed in studying platforms 'in action' which requires a detour through past STS approaches to media: firstly, because this is the analytic frame adopted by much of the existing literature, and secondly because it forces us to consider certain difficulties which arise when our access to the object of study is *mediated* in particular ways. I will first introduce the technique of controversy analysis, in particular, the version associated with ANT, and note that studies in this tradition seem to have historically downplayed the role of media in their analyses. However, a new set of literature at the intersection of STS and media proposes to analyse media, including online platforms, as technologies. I argue that this literature tends to inherit some of the difficulties of media studies and propose that these difficulties can be addressed through the STS concept of 'devices', which provides a richer understanding of technologies. This concept, however, also raises a possible tension between device-centred and controversy-centred analysis and I propose that the two need to be kept in tension.

The problem I raise in the next chapter (Chapter III) is that digital tools are entangled with platforms, and this poses problems for the conceptual programme outlined above. Latour and Venturini have argued that these online digital traces finally make possible methods which resolve tensions between quantitative and qualitative approaches and micro and macro scales. While this proposition is perfectly feasible in relation to web data, new online platforms tend to fragment the analysis. Digital Methods, which repurpose the relevance defining metrics of participatory platforms for social research, are the most promising; however they do not have a fully articulated qualitative equivalent. Also, I argue that to study controversies and not just platforms, it may be necessary to diverge from the data structures offered by these devices: moving past individual platforms to map various linkages between them and also moving past these quantitative traces to capture more than just the low-hanging fruit.

In the first empirical chapter (Chapter IV) I analyse the coverage of the Fukushima disaster on the English-language version of Wikipedia. Just as boundaries between science and media are constructed, but in practice crossed all the time, boundaries

within and between platforms and data formats must also be traversed in order to track controversies. One of these less visible phenomena is the references to external sites, which to a large extent drive the content in unfolding events. I map these references over time, consider how they are selected and how new boundaries are enforced between different types of sources.

Chapter V considers a newer 'social media' platform and how it contributes to anti-nuclear activism. I argue that while these platforms facilitate group definition and organising on the ground, they should also be thought of as directing interventions in purely online spaces – something I call digital demonstrations. However, social media platforms largely frame the success of these interventions in quantitative terms and this requires both analysing the performative effects of quantitative traces but also qualifying and contextualising them. I propose a visualisation which allows traditionally 'quantitative' and 'qualitative' traces to be read together. With this tool I show this quantitative phrasing of participation is slanted in favour of larger, more connected, pages.

In Chapter VI, I consider the diffusion of news stories on Twitter and how this requires rethinking the relationship between content and infrastructure. I develop a system of colour coded-strips to detect textual changes which both reframe the content of the story while simultaneously shifting the potential audience for it. Using the tool I identify three modes of diffusion which have consequences for how far information travels but also the contents of that information.

Finally, in the conclusion (Chapter VII), I sum up the potentials of these techniques for the study of public science controversies, where this work could be taken in the future and also some of its limitations. I also argue that the visualisations I develop can be used in an exploratory way, closer to the practices of natural sciences, rather being deployed primarily for the didactic communication of findings. This insight suggests that while digital tools can enhance online research, they should affect a slower process, through deeper engagement with their empirical object, which will always escape the devices that mediate it.

II. CONTROVERSY ANALYSIS AND MEDIA: DECENTRING THE OBJECT

'Facts travel by light-beams these days'
- Donna Haraway (Haraway, 1988: 597 note 1)

STS scholars have long concerned themselves with the possible impacts of new communication technologies on science and politics.[22] As Sheila Jasanoff recently remarked 'In the computer age, it is increasingly difficult to pin down with certainty the places where politically salient events originate, let alone to determine who controls the levers of power' (Jasanoff, 2004: 16). While the internet may furnish researchers with new equipment for researching controversies, it also appears to interfere with and complicate them in unpredictable ways.

In the last chapter, I proposed that online platforms may not only offer new sources of data to study science controversies, they also may shape their trajectories, and this warrants a more detailed study of these platforms 'in action'. I also noted some uncertainty about what role they might take in controversies: are they technologies for organising oppositional groups or potentially contesting scientific knowledge claims, or circulating alternate media articulations of the controversy? These possibilities are by no means separate, but in this chapter I am going to focus on the latter conceptualisation of platforms, in relation to broadcast media and journalism. This is important, firstly, because this is an analytic frame which much of the literature on online platforms adopts, and secondly because theorising the role of the media highlights some deeper ambiguities about what the object of controversy analysis is, specifically when our access to it is heavily *mediated* by particular entities like news media.

ANT is well equipped to deal with mediations generally but something about the particularities of the news media complicates matters. As I will describe, in the past, ANT-informed analysts of controversy and ethnographically inclined researchers in

[22] Haraway's remark in the epigraph was actually referring to satellites and communications *within* science, long before the internet as we know it.

PUS have often side-lined the role of the media in their analyses. There are several potential reasons for this: because media is seen as too 'light' a topic in relation to science, or because media cannot easily be traced in discreet networks. In any case this leads to certain deficiencies and tensions in ANT's understandings of media, which need to be addressed in order to understand the role of platforms in controversies, and in controversy analysis.

I should clarify, however, that most current work at this intersection of STS and media attempts to incorporate insights from STS into existing media frameworks, or carve up territories diplomatically between the two (Couldry, 2008; Wajcman and Jones, 2012). There is not scope in this project to provide a general theory of the media; my more modest goal in this chapter is to better attend to phenomena like media in controversy analysis – to attend to the capacities of particular devices and arrangements to articulate controversies and circulate these articulations.

One particularly helpful way of conceptualising media from within STS has been the concept of 'devices', which offers a more nuanced understanding of particular media technologies and how they format controversies. However, there is a potential with such approaches that the devices themselves become the focus of the study. Also, I argue that the object of controversies, if engaged with fully, raises questions about which devices are most central and how they interrelate. This leaves us with a tension: either we follow controversies and ignore the role of mediations or we study the mediations at the expense of the controversy and other media. I will propose that these two approaches need to be held in tension and the tension explored empirically. In the next chapter, armed with this better understanding of media technologies, I will focus in on the specific challenges of online platforms.

I will begin, however, by situating this study within the approach known as controversy analysis, in particular the strand associated with Actor-Network Theory (ANT) and studies of science 'in action' through qualitative, anthropologically inspired techniques.

01. CONTROVERSY ANALYSIS AND MEDIA

Controversy Analysis

The tradition of studying controversies (Pinch and Leuenberger, 2006) within science studies could be traced back to the study of 'priority disputes' (Merton, 1957), and the negative impacts of science and technology (Nelkin, 1971), but controversies were most famously introduced by the Edinburgh 'Strong Programme', as a lever to open up the content of scientific knowledge to social science scrutiny (Barnes, 1977; Bloor, 1976). Before this point, social science was only authorized to explain scientific *failures* in terms of the impact of 'social factors' or 'bias'. Accepted scientific *facts* on the other hand only needed to be explained by their correspondence to nature. The Strong Programme proposed a symmetry principle: that both scientific successes and failures are in some sense social phenomena and should be explained using the same resources. For example in Bloor's (Bloor, 1982) study of debates between Hobbes and Boyle over the corpuscular theory of matter, *both* positions in the debate were wrapped up with assumptions about politics and the proper organization of society. Pinch and Leuenberger (2006) helpfully clarify that scientific controversies, in this early usage, should be thought of in contrast to on-going disputes between, for example, religion and science, or epochal shifts like scientific revolutions.

Another version of controversy analysis, emerged from ethnographic studies of laboratories, and was foundational for the approach known as ANT (Latour, 2005). This approach proposed that controversies cannot be explained by 'the social' alone: the active role of nature and non-human technologies, such as laboratory equipment must be considered. The social thus becomes socio-technical. More broadly, ANT questions social science explanations based on macro-structures or culture (Latour, 2005), instead seeking to explain the stability of institutions through networks of associations – heterogeneous assemblages of both human and non-human entities, both referred to as 'actors' or 'actants'. ANT demands that researchers 'follow the actors', not deciding in advance which sorts of actors or what settings or strategies are most consequential for the stirring up and potential settling of the controversy (Latour, 1987). It also insists that the researcher not impose conceptual categories, like 'interests' (Barnes, 1977) on actors or sides of the debate, but rather, let them deploy their own categories. Finally, it requires not assuming in advance a stable reality or 'nature' external to the controversy – in the case of science, reality is often an *effect* of a controversy being closed. This constitutive uncertainty about the object of study, where conceptual questions become empirical questions, is what makes the ANT

version of controversy analysis unique and effective at dealing with complex or uncertain empirical objects.

Although ANT is, by design, not programmatic or codified (Law and Hassard, 1999), there are some features of this approach to controversies which inform the way I understand the slogan 'in action' and are relevant to the aims of this thesis. Firstly these studies make the link between specific *contents* of science (particular facts or knowledge) and the complex practices, infrastructures and materials through which they are produced. Secondly, making this link requires sustained qualitative analysis, either through the interpretation of texts or some form of participant observation, though quantitative techniques also may play a role. The end-results are *descriptions*, as opposed to causal explanations, of the process: how the controversy develops *over time* as opposed to static, synchronic slices. Finally, and this is the key point for this chapter, 'in action' studies require researchers to attend to the role of various mediations, classically scientific texts and inscription devices mediating nature, which sit between the research setting and the objects of study but also, non-scientific documents such as funding proposals and audit reports which negotiate relations between participants in the controversy. This means acknowledging that none of these devices or representations are perfectly faithful 'intermediaries', which circulate materials without altering them, they are all to an extent 'mediators', entities which transform or perform, what they represent (Latour, 2005).

ANT studies have classically focused on controversies over the settlement of 'matters of fact' between competing networks of scientists and their allies – revealing the constructedness of scientific representations was central to sociological critiques of scientific representations. But the same researchers have increasingly moved out of the laboratory (for example Latour, 1988) and turned to analysing 'matters of concern' (Latour, 2004; Latour and Weibel, 2005) — messy socio-technical entanglements which threaten society and must be accommodated within it. In this latter formulation, controversies are larger and more complex than mere matters of fact: not just epistemologies but ontologies are at stake.

This way of understanding public controversies is also informed by a pragmatist reading of democratic politics (Marres, 2005a, 2005b). In the 1920s, Journalist Walter Lippmann saw public opinion intervening when expert knowledge and traditional institutions fail to solve problems, but John Dewey saw 'publics 'as a very specific

entity which are defined by 'problems' which they were indirectly but personally affected by. Despite their different formulations of problems and very different solutions, Marres notes that in both accounts the public and the 'issue' (a term which roughly maps on to 'matters of concern' and 'controversies' emerge together).[23][24] Issues, in Marres' use of the term, however, have the connotation of public problems, which may or may not be science or knowledge focused (2015), and may not play out like public science controversies normally do, a distinction which will come in handy later in this chapter. Again, it is not within the scope of this thesis to further theorize how platforms are implicated in enrolling and shaping publics, the point I wish to take from ANT is that the question of which actors, or non-human objects, locations and technologies are relevant to matters of concern or issues is an empirical question, not determined in advance.

Actor-network theorists have also proposed that 'matters of concern' or 'issues' are not only increasingly common but also increasingly messy (Callon et al., 2001). Latour suggests this is because of the weakening of scientific legitimacy by the critical analyses of science offered by social scientists (Latour, 2004) but another possible explanation is that this has to do with the increasing availability of information about science made possible through media, specifically the internet. But while internet may help to distribute and complicate controversies, it also offers a means to chart them.

Mapping Controversies

Latour and other advocates of ANT have more recently turned their attention to analysing controversies digitally. Mapping Controversies is an inter-disciplinary research program, which provides tools to explore and visualise the complexities of 'scientific and technical debates'. This program, first of all, takes the form of a general methodology for analysing controversies. One influential formulation of this, which I will focus on, comes from Venturini (2010b) who describes the initial observation process as starting, as is often necessary, from a mass of competing *statements,* through which one can discover relevant *literatures*. From these bodies of references one can find human and non-human *actors* and then situate them in *networks* of actors. These

[23] When I refer to *publics*, I mean publics in this elusive, plural and materially entangled sense in contrast to '*the* public' as in the abstract object of political theory.
[24] The Lippmann / Dewey debate, interestingly, can be traced back to Lippmann's *Liberty and the News* (Lippmann, 1920) which described the inability of the public to process and make informed decisions about issues in a technical society based on the information made available by the news media. So the media and publicity was always at the heart of this question.

networks can then be consolidated into ideological camps or *cosmos,* the trajectory of which over time can be analysed as *cosmopolitics*, referencing the work of Isabelle Stengers (2010).[25] The end-result of these analyses are a series of interactive data visualisations, often but not exclusively networks, embedded in interactive websites with accompanying texts.

Venturini describes Mapping Controversies as ANT freed from certain 'conceptual complications' (Venturini, 2010b: 1 note 8). One of these I would argue is the provenance of their data. Researchers gather statements and literatures through search engines, websites, social media but also scientific citations or offline books or archives. As I just explained, classic ANT studies would need to consider the ways in which these materials *mediate* the controversy: for example how search engines rank the actors, or journalists represent key positions, but while these effects are acknowledged they are largely absent from the maps created within the Mapping Controversies rubric.[26] These media are mostly used *instrumentally* to disclose lists of actors and networks associated with the controversy, without their networks and chains of mediations being interrogated, or at least not as a central part of the methodology.

For Venturini, this streamlining is necessary because Mapping Controversies is first and foremost a pedagogical instrument designed for masters students across disciplines, and as a means of communicating findings to stakeholders and participants in controversies (Venturini, 2010a). While these cartographers of controversy do not claim to intervene as such (Venturini et al., 2015), these maps are intended to help participants and novices navigate these controversial landscapes.[27] Because of the communications requirements and the fact that the end results are graphic

[25] In one case study, for example, researchers mapped the controversies over the design of the 2012 London Olympic stadium, which including an interactive timeline, a networked map of the key actors scaled by their media attention over time and an interactive visualisation of different parties attachments to different aspects of the controversy: budget, legacy etc. (see also Yaneva, 2013).

[26] Venturini, for example, discusses the difference between search engines and the web, the web and the internet, the internet and the digital etc. (Venturini, 2010a) and how different digital monitoring devices may slant the study. His solution seems to be to multiply the monitoring devices and perspectives with multiple maps, as opposed to studying the monitoring devices themselves.

[27] The authors do however acknowledge that the act of describing is always in some sense also intervening (Hilgartner, 2000) but only to the extent that representations are taken up and implemented by actors subsequently. In contrast to academic studies of science controversies, Mapping Controversies is explicitly concerned with disseminating their work outside of the academy.

visualisations, the presentation of the controversy necessarily abstracts it from the various media or materials through which it was apprehended.

While this thesis builds on the work of Venturini and colleagues, it also seeks to extend and elaborate it by addressing a particular challenge: to map controversies while adding in some of the conceptual and methodological complications raised by the use of online platforms, and to develop some techniques to address these complications. As I argued in the previous chapter, online platforms not only mediate access to controversies, platforms may intervene in controversies as well. It is therefore important that we not only make use of the 'ready-made' products and traces of media disclosed by platforms to analyse the controversy. We must also consider how they are produced 'in action' to invoke Latour's phrase. While acknowledging that there may be limits to the small-scale technique of ANT for studying matters of concern, my gambit is that an ANT sensibility may help supplement the important work of mapping controversies with a richer understanding of the various technologies and media which stand between controversies and observers.

But before jumping to online platforms, I think a brief detour is necessary. Long before controversies went digital, one of the most important mediators of controversies has been the mainstream news media. Yet the media have been under studied and under-theorised from an ANT perspective. ANT specialised in representations and their circulation between 'centres of calculation' (Latour, 1987) but modes of representation outside of science seem to create conceptual difficulties. The reason it is important to consider these complications arising from the media is that many studies of online platforms, which adopt this analytic frame, seem to also inherit some of conceptual baggage which comes with it.

STS and Media

In the second edition of the *STS Handbook*, Bruce Lewenstein (1995b) provides a comprehensive survey of studies of the media within STS, including critiques of science journalism (Collins, 1987; Silverstone, 1985) studies about public opinion (Freudenberg and Rosa, 1984) as well as the role of the media in studies of controversies (Nelkin, 1971; Wilkins, 1993). Lewenstein makes two important points about the literature up to this point, which are relevant to the current discussion. Firstly, most of these authors do not pose the question of what *counts* as media: media

is assumed to be mostly journalism, as opposed to, say, science fiction, and it is also assumed that journalism largely comprises newspapers and television, as opposed to, say, radio. Although this is not the phrase he uses, these studies are therefore largely 'media-specific': they mostly attend to one medium at a time rather than together, in concert. In contrast, media anthropologists like Bausinger (1984) have located many different media interacting in 'ensembles' in domestic settings.

Secondly, Lewenstein laments the fact that this literature at the intersection of STS and media studies does not offer a sophisticated model of how the media works, which can match STS 'sociologically sophisticated' take on scientific knowledge. At the time, he argues, much work relied on linear 'sender-receiver' models often fragmented between studies of production, analyses of texts and studies of reception, as is often necessarily the case with media (Deacon and Fenton, 1999). Lewenstein also flags up the work of Stephen Hilgartner who similarly questions what he calls the 'dominant view of science communications' (Hilgartner, 1990) in which science gets the facts straight and then disseminates knowledge, in simplified form, to the media and then to the public.[28] Hilgartner contends that such distinctions between science and popular media are untenable: there is no discernable point where science stops and public facing journalism begins.[29] While Lewenstein does not advance an alternative model of communications here, he does suggest that it must take into account different sorts of media and interactions between audiences, journalists and scientists rather than just a one way flow of information.

These two related problems: the tendency to look at particular media as opposed to others, and the lack of sophisticated 'interactive' models of how information travels, I think need be overcome if we are to understand the role of online platforms in controversies 'in action'. But arguably such an understanding is still forthcoming even when we consider more recent work that investigates relations between science, technology and media. This may be, partly, because the front line of research into controversies has moved elsewhere.

---

[28] See also Chapter IV. Hilgartner contends that this is not how it works in practice – scientists obtain information from the media as well, and frequently produce information for different audiences – but only scientists are in a position to define what the correct flow of information is when it benefits them.

[29] Stephen Zehr, (1990) who examines the way scientific uncertainty is constructed, mobilised and managed, similarly shows *continuity* between the rhetorical and literary work of scientists in papers to similar devices in public hearings or more journalistic literature.

The Engaged Program

Lewenstein's critiques were echoed by other scholars in the emerging field of Public Understanding of Science (PUS) see (see Schäfer, 2012 for a survey of media literature in this field). Most notably, these scholars argued that the field to date operated on a deficit model: that authors were concerned with the scientific *illiteracy* of the public (Durant et al., 1989) the divergence between 'objective' scientific accounts and 'subjective' public accounts which takes scientific knowledge as a given and ignores the possible contributions of lay knowledge to the proceedings.

As Sismondo among others points out, in seeking to move past the deficit model of science communications, PUS scholars have shifted to an 'engaged program': promoting and / or studying sites of dialogue and participation — e.g. consensus conferences, public hearings etc. (Sismondo, 2008, also see; Ziman, 1991). Or in methodological terms, Michael and Irwin (Irwin and Michael, 2003) have called this an 'ethnographic turn' which aims to investigate *how* interactions play out between publics and scientists in the context of specific local settings. What is interesting about this literature is that researchers, who have rejected the deficit model, seem to avoid detailed analysis of the media itself, despite the fact that media still may play an important, or especially important, role in their accounts.[30]

To offer one prototypical example, in *Acting in an Uncertain World,* Callon, Lascoumes and Barthe (Callon et al., 2001) discuss the management of sociotechnical controversies through what they term 'hybrid forums' — clumsy, yet ultimately necessary, gatherings of heterogeneous actors which engage in collective experimentation and learning, where science (secluded research) confronts novel forms of lay-knowledge (research in the wild). But as these two knowledge systems clash, the media is frequently lurking in the background: disseminating information and promoting various positions. Callon, Lascoumes and Barthe assert that media coverage generally benefits hybrid forums, by making positions in the debate perceptible to the other parties. But in their account, media often seems to be more of a resource to be tapped into, or an ether through which the controversy flows, rather than an alternate channel in which the controversy plays out.

[30] Or to give a more empirically focused example: local newspapers were, one of the ways in which the excluded farmers studied by Wynne learned about the controversy and formed opinions, but the newspaper's role was not fully explored (Wynne 1992).

So why might it be important to have a more nuanced understanding of the role of media in controversies? There are multiple ways to approach this, but I will focus on the way PUS thinkers use the concept of 'framing' (Bateson, 2000; Goffman, 1974).[31] 'Framing', broadly refers to the way the controversy is defined, who it is made relevant to and on what terms (Wynne, 2005). Engaged program studies often reveal how the settings of science / public interactions tend to favour the expert or science framing of the controversy at the expense of alternative framings.[32] As discussed in the last chapter, the fact that the Windscale Inquiry was focused on safety issues (as opposed to economics or nuclear weapons) was instrumental to the outcome and also determined what counted as legitimate participation (Wynne 2011). However, much work in PUS is concerned to demonstrate that this definition or framing takes place *outside* and *in advance* of the settings of participation, as Marres highlights (2005a). Following Jasanoff, she argues that in order to follow the controversy or locate the site of politics, PUS might need to look *outside* of these institutionally sanctioned settings of engagement.[33]

It should be clear that the media (and specifically online media) is one such setting or channel though which framings of controversies are defined outside of or in advance of these more stage-managed interactions between institutions and publics. For example, Sheila Jasanoff in *Designs on Nature* (2005) analyses how scientists, politicians and activists fought over the GM foods crisis in Britain. Her book is all about 'collectively held' cognitive framings of issues, particularly national framings, but she only implies that the media might be one way in which frames are inculcated.[34] Jasanoff's focus is mainly on the policy dimensions of the conflict but she does discuss two media focused

[31] In its original usage, framing describes a kind of meta-communication through which a particular situation or event becomes intelligible – the signals that make other signals mean something – for example smiles or winks or exaggerated gestures which, in context, signifies that a fight is a play-fight (Bateson, 2000; Goffman, 1974).

[32] This is especially true of studies of public participation initiatives, such as consensus conferences or science exhibitions. Irwin (Irwin, 2001) discusses how institutions largely determine the framing of controversies and, thus set the terms by which 'the public' are allowed to participate, including normative understandings of citizenship (Wynne, 2008). There is of course scepticism within PUS over government run participation projects such as GM Nation which may be just token gestures toward the public used to legitimate state actions (Thorpe and Gregory, 2010).

[33] Notably, Marres suggests that online media, but also less obvious settings like the domestic sphere, offer ways of studying issues outside of institutional settings, before they become fixed (Marres, 2012a: 147)

[34] Jasanoff earlier cites Benedict Anderson's (1991) study of how national identities are constructed, largely through books and newspapers, which would provide one account of national framings.

events —Prince Charles' editorial in a British Tabloid and Greenpeace's 'truckload of beans' protest, which 'played directly to the media'.[35] These two interventions, one by a powerful actor and one by media-savvy activists, circumvented the normal policy-discussion process and altered the content of the controversy. Yet Jasanoff and similar scholars do not provide a detailed account of these media-inflected shifts.

So as the engaged programme developed more sophisticated understandings of science and lay-person interactions, they have not necessarily developed more sophisticated understandings of media. This is despite the fact that media may represent one space in which frames, which may shape encounters between institutions and publics or experts and laypeople, are developed. In the next section I will briefly turn to literature coming more from the media angle in search of a sophisticated understanding of media, which better understands the specificity of media practice: how it mediates and circulates controversies. The additional problem this highlights is that, until more recently (Gillespie et al., 2014), these studies tend to analyse frames in terms of discursive content, *or* study the material arrangements through which they are produced, but not the two together.

02. MEDIA AND MEDIA TECHNOLOGIES

Media Frames

Before jumping to STS informed studies of online media, it should be acknowledged that framing, as well as being a topic in sociology and political theory, has been widely discussed in media studies. Gitlin defines frames as the journalist's '…principles of selection, emphasis, and presentation composed of little tacit theories about what exists, what happens, and what matters.'(Gitlin, 1980: 6). Media frames coalesce into systematic ways of defining an issue *or groups* which, in his particular example, can determine the success or failure of social movements. So frames are not just descriptions of a controversy or event but they also they shape the topics they describe and potentially limit the range of debate (D'Angelo, 2002). They are, to anticipate a term used later in this chapter, *performative*.

[35] This was where activists dumped genetically modified beans in front of the Prime Minister's residence at no. 10 Downing Street.

This work also draws attention to the reusability of frames (see Kitzinger, 2000 on 'media templates' the use of past events to articulate new ones).[36] Frames also include within them a degree of controversy, that is positions and debates between actors who fit the frame, which obscures or occludes actors and positions outside it (Gamson and Modigliani, 1989). D'Angelo (2002) details three overlapping strains of media framing research: 'cognitive', 'critical' and 'constructivist'. The 'cognitive' strain of framing research looks at how frames in media impact the interpretation or evaluation of the issue by audiences (e.g. - Price, Tewksbury & Powers, 1997)(e.g. - Price et al., 1997), normally through surveys and focus groups. The 'critical' strain investigates how 'elite' framings — those that benefit the position of powerful actors — are favoured in the media by the professional routines, norms and values of journalists (Entman and Rojecki, 1993; Gans, 1979; Glasgow University Media Group, 1976) either through discourse analyses of texts (Fairclough, 1995; Van Dijk, 1993) and / or ethnographies of newsrooms (Tuchman, 1978). The 'constructivist' strain, which is most compatible with the PUS usage, examines how competing frames participate in the definition of social reality (Gamson and Modigliani, 1989; Molotch and Lester, 1974).

These studies, particularly from the latter two strains[37] could help enrich the PUS and STS understandings of the role of media in controversies, because they reveal the process and politics behind media products: in this case particular frames. But there are limits to this use of framing from an ANT-informed perspective on controversies. Framing studies based on ethnographies of newsrooms, which link everyday practices and routines to the production of media texts and 'reality' resemble 'in action' studies of laboratories (see Schlecker and Hirsch, 2001 for an explicit comparison) but they often identify frames as mapping onto ideological or class positions, or at the very least 'newsroom cultures' and thus rely on social explanations. Secondly, in each of the different strains, frames are cast as relatively stable entities which exist outside of individual articulations of them. As Marres has pointed out, the fixity or taken-for-grantedness of frames is precisely problematized by controversies which elude attempts to contain them (Marres, 2005b). As Marres also notes, whether in PUS or media studies, frames generally refer to discursive or more narrowly textual

[36] For example, Gamson and Modigliani describe how for the first 20 years of the nuclear power industry, most nuclear-related stories were interpreted using a 'progress' frame (nuclear power was an advance on coal), but with the events at Three Mile Island, the dominant media shifted to a 'technology gone wild' frame.

[37] Also, framing studies of the cognitive strain, entail claiming access to the mental states of research subjects, while an ANT account would see these states as *produced* through social science methods. See next chapter (Savage, 2010).

contributions. It would be important to also consider how different communications technologies and non-humans are implicated in daily news work and the fact that the world outside is not merely constructed through representations but also influences them.[38]

So perhaps one reason why ANT is reluctant to study media in depth in this way, is because object-oriented understandings of politics are premised on a critique of discursive understandings of politics (Marres, 2005a), which is how media are often approached. But ANT-influenced researchers are not against texts, in fact the literary analysis of scientific texts was central to its foundation, but it conceives of texts in different ways, as material and materially entangled (Law, 1986). However, as I will explain in the next section, more recent work at the intersection of media and STS moves past constructivist approaches to study the materiality of media, but in the process seems to lose this focus on representations all together.

So, although framing as a concept needs work, I have included this discussion to show how STS must understand the specificity of how media mediate and shape controversies through, perhaps more partial or fragile definitions, something which I will develop further in the next chapter.

The Material Turn in Media

There is, today, a rapidly expanding body of work at the intersection of STS and media which addresses itself to the historical neglect of materiality in media, though attempts at such a dialogue have been going on for some time (Couldry, 2008; Livingstone and Lievrouw, 2009; Oswell, 2002; Silverstone and Hirsch, 1992). Leah Lievrouw, in an important edited volume *Media Technologies* (2014), argues that this is part of the swing of the pendulum: while early media and communications could be criticized as technologically determinist, the idea that technologies are drivers of shifts in society and culture (for example, the work of Harold Innes), it has now swung too far towards social constructivism – the idea that society or culture shapes uses and meanings of

[38] In my opinion, the closest study to viewing the media 'in action' is Gaye Tuchman's *Making News* (1970). Tuchman describes how the temporal set up of the news day, and organisational typefications: 'spot news' 'business news' 'politics' etc. all shape how information is processed and what becomes 'newsworthy'. Tuchman applies these ideas to the specific instance of the women's movement and how the timings of their press conferences and the emergence of 'women's sections', even internal politics of the newsroom, fundamentally shaped how the issue was covered in the long term.

media technology. Today the stated goal of much work in this area is to describe the 'co-production' or 'mutual shaping' of society and technology, inspired by STS, but she argues that most current work still remains 'tilted' toward the social side. While this is a fair assessment, and the role of non-humans in media must be pushed further, in this section I want to point out that recent moves towards materiality have actually been at the expense of content as such, or at least maintained their analytical separation.

One of the most influential studies of media and materiality has been Roger Silverstone's celebrated *Television and Everyday Life* (Silverstone, 1994). Silverstone, who embraced social construction of technology, but was somewhat ambivalent about ANT (see Couldry, 2008) described media as 'doubly articulated': as both a material, technical object of consumption, situated in the household, and as a purveyor of content, and, crucially, the interrelation of the two. As Sonia Livingstone more recently pointed out, however, followers of Silverstone have for various reasons tended to study *either* the materiality *or* the content, a tradition which continues, she claims, in internet research (Livingstone, 2007).

A more recent development has been ANT-informed ethnographies of the newsroom, presumably because of its obvious affinities with the small group-setting of a laboratory. Most of these newsroom studies are concerned with innovation and the advent of new technologies (including participatory media) (Boczkowski, 2005; Plesner, 2009; Schmitz Weiss and Domingo, 2010).[39] Yet unlike the ANT studies of science 'in action' they generally do not relate technologies to the production of *particular* facts or content.[40] Emma Hemmingway's study *Into the Newsroom* (2008) focuses, in contrast, on how journalists' *everyday* entanglements with technologies, impact the *content* of the news in the way 'inscription devices' (Latour & Woolgar, 1979) impact the finished scientific article. However, Hemmingway focuses on individual news 'facts', which I argued earlier may not be the most appropriate unit of analysis.

---

[39] As van Loon points out, there is nothing necessarily 'ANT' about interactions between technology and humans – this threatens to be merely material culture – their interdependence should be presumed from the onset. Also there is nothing special about *new* technology which makes ANT finally applicable to media. (Van Loon, 2011).

[40] Plesner (Plesner, 2009) however analyses the role of ICTs and other technologies in news work viewed through the lens of actual media articles (and supplementary interviews). But instead of focusing on ICTs as a separate entity impacting news work, she examines how they have been naturalised – successfully 'translated'. But Plesner also does not discuss how these entanglements impact the form and content of the texts themselves – only that they were successfully published.

Yet, attending to materiality does not necessarily entail observing an offline setting.[41] Another focus for these new intersections has been particular online platforms or 'formats' like blogs (Siles, 2011) and the materiality of certain online technologies. It may seem like a counter-intuitive proposition to study the materiality of an object which is so generally thought of as dematerialised, or 'virtual', but as Mathew Fuller (2008) argues in relation to software, thinking of the material is not necessarily about the circuit boards and blinking pixels but also the code and software which permeate much of social life today. This is the premise of 'software studies' on which some work at this intersection draws. So investigating the materiality of online media does not necessitate handling artefacts, like computers, laptops or smart phones, or the infrastructure which makes the internet possible, but might include the little non-human bots and scripts and design features which are embedded in social life today (Fuller, 2008). There are, of course, costs to abandoning a fieldwork setting in terms of empirical detail and contextual understanding but with the object of matters of concern, which tend to transcend any particular setting, this may be a necessary move.[42] Also, since it has long been acknowledged that online and offline worlds are entwined, not separate, the online as Rogers argues, may be a feasible site from which to 'ground' studies about collective life (Rogers, 2009). This does not necessarily mean rejecting offline fieldwork or in-person interviews, merely that the online or particular technologies may provide ways of locating settings, actors and objects which can be followed up with other methods.

A related strain of work could be characterized as platform studies or studies of 'platform-specificity' (Elmer, 2006; Rogers, 2013b). Although this is a very broad and varied field (see Helmond, 2015 for an extensive review), these studies generally focus on platform-specific objects (links, hyperlinks, tags, results, hashtags etc.). To give an example, Gillespie considers how algorithms, these increasingly ubiquitous bits of code and scripts, '…manage our interactions on social networking sites, highlighting the news of one friend while excluding another's' (Gillespie, 2014: 167). Every day Twitter's 'trending' algorithm identifies top terms and hashtags for a particular region,

[41] It also seems counterproductive to limit ones self to an offline site like the newsroom when new media literature argues that this is no longer the only site through which information flows and decisions are made (Bruns, 2005)

[42] Many STS scholars have however, successfully extended interactionist observations beyond traditional fieldwork settings, taking into account, in the case of remote interactions via computer networks, the various equipment or 'scoping devices' which mediate interactions (Cetina, 2009).

though it is unclear how this is calculated or what 'trending' means (frequency? popularity? acceleration?). Gillespie describes how users have revolted against Twitter and other algorithmically based services at the suggestion that un-marked advertising, paid-for-content, would be featured in results lists, which demonstrates how successfully these algorithms have become imbued with a sense of objectivity. Much like the original STS interventions into science, Gillespie is using an STS-inspired approach to reveal some of the subtle politics of technologies, which are otherwise presented as neutral or objective. Yet he is careful to not turn this into a techno-determinist account: he also considers the discursive and rhetorical constructions of algorithms as impartial and also the assumptions of human programmers and recursive contributions of users who are attempting to game or anticipate algorithms. Also the interest for Gillespe lies in how these algorithms impact or influence the *content* of these platforms, though he does not consider particular contents in his brief account.

The point of this brief survey is that with a few exceptions there is often a tension between a focus on materiality and technology and a focus on the contents of media, which I argued earlier was necessary to understand media contributions to controversies. Bockzowski and Siles in the same volume as Gillespie (2014) affirm that, beyond the techno-determinism / social constructivist dynamic, the recurring problem for STS and media studies is a twin set of dichotomies: the balance between studies devoted to *content* and *materiality* and *production* and *consumption*.[43] In a similar schema, Wajcman and Jones (2012), drawing on successive theorisations of media by Hall, Silverstone and Du Gay, see the media as comprising two interrelated 'circuits': a) the production / consumption of *material* artifacts like TVs and computers and b) the encoding / reception of media messages they carry. They then argue for a unification of these circuits in which ANT is assigned the analytic weight of the material production / consumption circuit.

Both the twin axes and twin circuits proposals reproduce the idea that ANT specializes in the material, technological artifacts and non-human nature, but is not conducive to studying content. But as Lucy Suchman (2014) argues in her response to Boczkowski and Siles, in any given empirical case study, the content-material distinction should not

[43] As Boczowski notes, these conceptual divides are just as much a practical methodological problem of different types of data and methods (ethnographic or content analysis), and different department specialisms. I will deal with these more methodological issues in the next chapter.

hold anyway, at least for ANT inspired approaches.[44] In the next section I want show why this is the case with the approach known as Digital Methods and the concept of 'devices'.

I will not presume to resolve these tensions or bridge these divides, which present very real challenges for media scholars trying to articulate a division of labour within their field. Also, some of these concerns for media, such as the political economy of media and the symbolic role of technologies and messages, are best left to other approaches (Couldry, 2008; van Dijck, 2013).[45] I will instead propose that for the purposes of controversy analysis, these distinctions become an empirical problem rather than a conceptual one.

03. DEVICE OR CONTROVERSY?

In this final section, I am going to attempt to show that when researchers consider not just materiality or co-production but some of the other equipment of ANT-inspired controversy analysis, namely the concepts of 'devices' and 'matters of concern', these conceptual divides (content / materiality and production / reception) become less troubling. For the time being I will address 'content' as mostly discursive, textual products of media, which is often what is meant within media debates, but then I will consider another understanding of content as 'substantive topic' being discussed.

First I will discuss the STS concept of devices, which has been successfully deployed by Digital Methods, among others, in relation to platforms, but then I will propose that such an approach to platforms could be contrasted with approach, centred on controversies.

The Device-Centred Perspective

One way to complicate and enrich the way that current approaches talk about media and specifically online *platforms* is through the ANT use of the term 'devices' which

[44] Similarly, ANT studies should, ideally, trace the circulation (and translation) of materials across domains of production and reception.
[45] Van Dijck for example uses ANT to talk about technological development and 'closure' but interestingly combines this with a more Castells influenced account of economy and consumption.

builds on a co-construction or socio-technical understanding of technologies. Devices were crucial in the early critiques of scientific representations (Latour and Woolgar 1979) and more recently Marres, (2012) in relation to studies of participation, advocates a 'device centred' perspective: that is focusing on the role of particular objects, technologies or settings in enacting participation. Devices help with understanding the media and particularly platforms for three reasons: 1) they have indeterminate boundaries 2) they are performative: they both represent and enact forms of sociality 3) they question the textual-material divide.

Devices could be traced back within STS to 'inscription devices' or the laboratory equipment which turned unwieldy nature into manageable traces and numbers and were fundamental to the strength of particular facts.[46] John Law and Evelyn Ruppert recently described devices as 'patterned teleological arrangements' (2012). They are heterogeneous assemblages of humans, non-human technologies, texts, institutions, nature etc. So rather than a dialectic of social and technological shaping, this interrelation is presumed from the onset. What Law and Ruppert mean by 'teleological' is that devices do things in the world but they are not reducible to their intended designs. Devices are, as Michel Callon, puts it in relation to economic instruments, 'performative' (Callon 1995).[47] Rather than merely representing the world, they enact certain arrangements, but 'performance' entails unintended consequences, as its success or failure depends on the cooperation and coordination of heterogeneous actors.[48]

Callon's use of device also builds on Goffman's use of the concept frame, just like the more discursive-focused understanding encountered earlier. In his example of economic markets (Callon, 1998), economic transactions are able to take place because certain associations have been temporarily, imperfectly, disentangled from each other through the use of 'frames', to use one of Goffman's original metaphors, like theatre

[46] Although there is not space to discuss it here, device has an even longer history which could be linked to the apparatus [dispositif] of Foucault and assemblages of Deleuze which have different analytic potentials (Law and Ruppert, 2013). I am here using device in a general sense to problematise more monolithic readings of technology, but it could be argued that the particular platform/devices in the empirical chapters demand these different but related understandings.

[47] Callon's argument is that abstract economic theory (in the form of text books and articles) is used to justify technologies and arrangements, which to some extent make economic theory applicable in the real world. So rationality may not be a natural characteristic of humans but it can be approximated when individuals are equipped with technologies allowing the instantaneous comparison of prices and models for making decisions.

[48] Do they follow these 'scripts' or produce their own 'anti-scripts' (Latour 1992).

performers and their audience are kept separate by the theatrical stage and curtains.[49] But this disentangling of countless messy entanglements and their effects, which Callon refers to as 'overflows,' is always partial and incomplete because what happens in the frame often requires access to these overflows as resources.

When Gillespie critiques the self presentation of 'platforms' (2010) as offering a neutral stage for debate, this could be though of as the 'device frame', the intended or implicit *modus operandi* of a technical arrangement which will never be fully realised in any given scenario. What device adds to the study of platforms, and Gillespie is certainly aware of this, is the understanding that platforms are nested in and depend on larger socio-technical systems: legal institutions, active audiences, other competing platforms etc. This means that, following Callon's (1998) suggestion, one can look for controversies and contests over device frames and this may involve tracking overflows: not just looking at frontstage phenomena within the frame but also backstage processes which fall outside it.

So although I will occasionally refer to platforms as devices, and refer to their *modus operandi* or device frame, it should be understood that these devices are not coterminous with the platform's nominal boundaries, they do not stop at the web domain (Facebook.com). Platforms are also composed of devices, particular scripts and objects which may cut across them. Methodologically focusing on these objects and components will help to get beyond the frame, beyond the device's self-presentation. But as Law and Ruppert (2013) observe, while following a device's chain of associations – the overflows – is a potentially endless task. The boundaries of a device-centred study are at least partially defined empirically in the sense that devices readily offer up certain empirical materials, certain avenues and contours to follow.[50] Some of these boundaries are also practical, and methodological ones (discussed in the next chapter) but they are also according to Law and Ruppert (2013), political: determined in part by the researcher's agenda and questions. Revealing the extension of the device's overspills, either as unintended effects or obscured inputs could help one make sense of the workings of power or reveal the contingency of supposedly neutral technologies.

---

[49] Callon emphasizes that 'performative' is different from 'prescriptive' which is a type of performance in relatively closed situations where roles and actions can be repeated indefinitely. (Callon 2006)

[50] Such as trying to keep media separate from advertising or obscuring shady sources, or distinguishing science from breaking news desks.

This raises the following question: in following Gillespe, trying to attend to the 'politics of platforms', does this involve tracing platforms back to content producers in offices or users and active audience members perched at their laptops? Another advantage of studying platforms as devices is that we can to some extent empiricise the producer / audience distinction. Recent work on 'devices of the public' or 'participatory devices' (Marres, 2012a; Marres and Lezaun, 2011), deals with the role of objects, technologies and settings in materialising or making visible publics and participation.[51] For Lezaun, a focus group is enabled by a specific material setting – a sterile white room with a two way mirror (Lezaun, 2007) which produces novel forms of behaviour in its subjects. Such devices and settings also enact particular forms of inclusion and exclusion. One could make an, albeit clumsy, analogy between particular venues of social media platforms and a circumscribed setting like a consensus conference, forum or a focus group and describe how audiences, publics or more specifically participation is counted and materialised through platforms. Just as public hearings set the terms on which participation could happen, the architecture of online platforms favour certain types of actors, types of evidence and practices at the expense of others. For example: rankings and search results may privilege more popular or networked users; forms of automated moderating may censor certain kinds of speech; or routines and customs may benefit more experienced actors. Platforms can thus be evaluated in terms of what modes of participation they enable.[52]

So rather than start from audiences or content producers as an object, one may instead follow the device and ask how actors, participants, publics or audiences are made visible as such by these means. This would be to build on studies, of television for example, that look at audiences as rhetorically and materially constructed through various technologies such as set monitors and home diaries (Ang, 1992).[53] But, as

[51] This work challenges past models of participation from political theory which are premised on normative models of public deliberation as primarily *discursive* but also post-Foucaultian approaches to politics which study the material aspects of governance but which also see the material as 'under-articulated' or even purposefully obscured from view. In trying to bring texts or content back into the picture I do not want to reproduce a discursive or dialogic idea of politics but analyse the work of texts as material and in relation to the material.

[52] Studies of participatory devices do not make assumptions about the form that participations or materialised publics take, this is an effect of the device. This is perfect for social media in which publics may not be, in the classic Deweyan sense, as materially implicated in the controversy, but may be hobbyists or conspiracy theorists, with only a passing interest but 'enrolled' through the device.

[53] It has been convincingly argued in media sociology that studies which look at the discursive construction of the audience should also mobilise ethnographic approaches (Ang, 1991) or

Livingstone remarks, in the case of TV, the elusive audience reception of content was largely tucked away 'in people's heads', but on the internet it becomes more readily visible, through comments, posts, blogs, likes and other technologies.[54] Similarly, but perhaps less obviously, the internet also makes available limited traces of the production of content, such as timestamps, past versions of articles and hyperlinks to press releases and sources. But these traces of audiences and traces of production processes, should not be taken for granted, but investigated as topics.[55] In fact, these platforms which allow a wider range of users to produce content, problematise distinctions between audiences and producers (as well as mainstream and alternative media) *empirically*.

But these platforms do not just enact particular forms of sociality and participation they also produce certain sorts of mostly-textual products, which is what I was gesturing to earlier with the concept of media frames. It should be clear that this understanding of media frames introduced earlier is not unrelated to what I referred to as device frames just now. Both have their roots in Goffman and both in some sense describe how a situation or event is made intelligible and manageable to participants. They are both, as discussed earlier, *performative* in the sense that they can shape the controversy: influencing politicians and informing the terms of debate in public hearings.

The media definition of framing can be seen in Callon's terms as a temporary clearing of a controversy: a statement of the key players or an assessment of what is relevant, which works by disentangling or excluding certain actors, terms and sub-issues at the expense of a messier reality. So if a commenter on a web forum were to frame a nuclear power plant as 'low carbon' source of energy, they would do so by severing it from the

---

more forcefully they need to consider how audiences and institutions are interrelated as part of socio-technical assemblages (Oswell, 2002). Although this study must remain silent on the audience as such, I will pair performative analyses of metrics with different sorts of qualitative analysis, though not ethnography, strictly speaking.

[54] She explains "..new media use is at least partially visible, for people must, necessarily, interact overtly (through selecting, clicking, scrolling and typing,) thereby coinciding in an auditable manner the symbolic and the material…"(Livingstone, 2007: 4)

[55] Rather than dealing with audiences explicitly as a topic, I will instead consider how 'potential readership' 'engagement' etc. are mobilised by participants through the proxy of rankings and digital traces and the drive to share. I will also consider how user contributions are labelled as experts, laypeople or members of the public. This can even lead researchers to identify potential producers and users of content to use in follow up interviews. In this way the online may make a better site to ground the study than any particular offline location (Rogers 2009). However, as Gillespe notes, in relation to algorithms: the idea of a backstage, that someone is actually in charge, itself is performed.

process of uranium mining and refining as well as the construction and lengthy dismantling process, all of which are carbon intensive, pushing these associations out of the frame. The plant becomes unmoored from its many real-world networks so that it can be abstractly compared, favourably, with other energy sources. Such a framing of nuclear might, of course, contribute to getting the plant being successfully built.

But these media frames, which are mostly understood to be discursive or textual products are also produced through socio-technical arrangements such as the participatory media platforms described above – as well as being influenced by what happens in other settings. The above hypothetical presentation of nuclear on a forum would be a result of conversations enabled by online platforms but also dependent on existing framings and on the particularities of the power plant itself – which will constrain possible attempts to describe it. In other words it is not the web commenter alone who 'frames' the nuclear plant but a distributed accomplishment.

What we have to abandon about media frames is the idea that they exist outside of particular textual/material instantiations of them. Rather than free-floating frames held in collective consciousness or individual minds, inculcated through social structures or immanent in corpuses of texts, it is better to think of these media frames in more actor-network terms as continually emergent, a *process* of framing which is unstable and incomplete, rather than a result. From now on I will use 'articulations', Marres' preferred term, to refer to the particular capacity of media to format controversies, through textual or material means.[56]

One approach which successfully bridges the content materiality divide with devices is Richard Roger's Digital Methods research programme, which attempts to repurpose the dominant devices of the web to map wider socio-cultural trends and particularly controversies (Rogers, 2013b). They, for example, analyse how keywords around particular issues shift over time in collections of pages, or how images embedded in Wikipedia articles differ between national contexts. Roger's work shows the relationship between particular contents and the politics of platform features, how platforms *format* issues or controversies. This work is largely interested in mapping dynamics over time, in a parallel way to Mapping Controversies, as opposed to more

[56] Articulations, of course has an equally complicated history both in media (Hall et al., 1980) and political theory (Laclau and Mouffe, 2001) but I will restrict my usage here to Marres's meaning in terms of issues taking shape.

granular, qualitative 'in action' studies, though in the next chapter I will propose how they could be used together.

However while 'device' helps bridge, or problematise the content / material distinction raised above, it also poses other problems. What I think is *really* at stake in the above distinctions for STS researchers is the tension between studying substantive objects in the world through media and studying media devices and platforms themselves (Marres and Weltevrede, 2013). Are researchers interested in devices framing / enacting online participation or framing / enacting offline controversies? Both of these involve textual and material processes but they entail very different sorts of analyses.

What I have described above is a device centred approach to media. Though I have used the concept of devices to problematise the text / materiality divide and show the contested boundaries of devices with associations stretching far beyond them, this still is to define the study through particular technologies rather than particular substantive topics or issues. While this is a viable way to conduct a study of media technologies, there is a danger, from the perspective of cartographers of controversies particular platforms are not in and of themselves are that interesting. They are sometimes invested with 'normative political capacities', as in situations like the Arab Spring (Meraz and Papacharissi, 2013) arguably in a similar way as explicitly participatory devices or carbon accounting meters (Marres, 2012a) but this is not guaranteed. They may be, as Morozov (Morozov, 2012) contends, built for entertainment, not for politics. Platforms might need particular controversies to animate them. This is the possible limitation of platform specificity: if we get so wrapped up in the scripts, bots and customs of platforms, we risk losing track of why they were interesting in relation to controversies in the first place. The other danger, though, is that we cannot know *a priori* which devices are most important to the controversy, or, from the vantage of particular devices, how they interrelate.

In the next section, I will suggest that when studies focus on controversies as an object they have certain benefits in relation to the problems identified by Lewenstein earlier in relation to the circulation and interconnection of media.

Controversy-Centred

So what would a controversy-centred study look like? Firstly it is worth noting that in the sense offered earlier, matters of concern have much overlap with devices. Ruppert and Law (2013), appropriately enough, give the example of the Fukushima Nuclear Power Plant as a device which: produces electricity, provides jobs, but as the site of a disaster also may enable certain forms of politics. The subtle or not so subtle difference is that devices such as the Fukushima plant become matters of concern when they shift from *objects* to *things*: from taken for granted bits of infrastructure to broken down, indeterminate problems which gather heterogeneous actors around them (Latour, 2004).[57]

This is of course a spectrum, both online forums and stricken nuclear power plants can be controversial and assemble participants but what concerns me here is whether the sleek, working device or the messy matter of concern is taken as the centre of the study. Does the researcher use particular media as devices and locate controversies over their participatory or democratic capacities or does the researcher choose controversies in order to investigate which sorts of actors, technologies and settings become implicated? If one is 'following the actors', then where the study stops is relatively open but this depends very much on where the study *starts*.

One controversy-centred example which is not confined to particular media is Lewenstein's (1995a) analysis of the 'cold fusion scandal'. Scientists Pons and Fleischmann held a press conference about their apparently successful demonstration of cold fusion to gain the attention of funding bodies and circumvent the normally slow peer-review process. After the initial upheaval this created, opponents of their claims held another press conference in order to ritually excommunicate them (see also Gieryn, 1999). Lewenstein sees the media as largely responsible for the shape and unfolding of the controversy, from the initial informational chaos to the consolidation of for / against positions and some of the key turning points. But media not only destabilised and accelerated the controversy, the actions of journalists and editors also affected the *content* of the controversy, introducing many of the economic and social terms in which the scandal was articulated.

This study is interesting for several reasons. First, Lewenstein shows the media-science relationship as bi-directional – media does not simply 'translate' scientific knowledge –

[57] Although I think it is important to note that the working nuclear plant and the broken nuclear plant are no longer the same entity

science responds to and orients itself to the media.[58] Second, Lewenstein discusses not only mainstream news, but also a diverse array of less obvious 'media': policy reports, email, audio recordings, pre-prints of scientific papers, faxes. For this reason, he argues that a 'circuit' of media communications hardly describes the complexity of this 'web' of communication channels and back channels. The particular arrangement of media technologies, institutions and actors only comes into relief through the *contingency* of a particular controversy, here the cold fusion scandal, which importantly is not just *covered* by media but includes the media within it.[59] Because Lewenstein focuses on a particular controversy, the question of which media is involved or what constitutes media in the first place becomes an empirical question which might be lost if he started with a media-specific analysis.

Questioning relationships between types of media and also sender-receiver, or circuit, models of communication is especially important in relation to online platforms. As Van Dijck makes clear in her overview of social media (2013), these platforms are particularly closely intertwined, in terms of algorithms and business models and cultures but also in terms of content – that is media content passes between them, and I would argue, passes between them and other entities like the 'mainstream' news.

Another example of a controversy-centred study, involving the media, would be Andrew Barry's account of a protest over a road construction cutting through a nature reserve (Barry, 2001).[60] In contrast to many PUS studies, Barry addresses the role of the media directly, explicitly making the link between science and news media as fact building endeavours. He discusses how the 'reality' of what happened during the protest was a strategic negotiation between activists, journalists and larger organisations like Friends of the Earth. Barry forcefully argues that their contributions are not reducible, as critical media studies account would have it, to either ideology or representations of external truth.

---

[58] But Gregory and Miller argue that it is important to take the media's role seriously because sometimes the media sets an agenda for the scientific community as a whole (Gregory and Miller, 1998). It is also important to note, as Gregory and Miller do, that not all scientific controversies appear in the artificially circumscribed 'science section' of newspapers.

[59] This study would also be an example of how the impact of the media could be demonstrated without citing particular media frames and 'presuming effects'. The effects are here judged by the scientists actions in response, rather than presumed shifts in opinion.

[60] Barry makes the point that some STS and post-Foucaultian accounts want to see politics everywhere but actually political action is rather confined and localised. It in fact takes much work for politics to become visible.

Barry points to a more complex picture of the media, not merely powerful corporate media against powerless activist media, but recognising independent journalists and activist organisations also have the bargaining chips to contest dominant representations. Controversy analysis allows Barry to problematize the relationship between mainstream and alternative news, audiences, sources and journalists *empirically*. It also allows him to examine interactions between the media and the objects being described by the media, circumventing social constructivist accounts. This is an excellent solution, which in contrast to most PUS studies manages to study the media 'in action' – that is both the products and the process, yet it is actually quite rare today for controversies to be localized in convenient offline settings (Latour, 2004) or indeed online settings or platforms.

However, the main concern with a controversy centred approach is that, unless the researcher is there on the ground, and arguably even in that case, our access to matters of concern will *always* be mediated, by which I mean 'translated', by media or other types of devices. Researchers can on one hand use the controversy to undermine media-centric analyses but they also rely on the media to *access* the controversy in the first place.

Decentring Devices

To bring this discussion back to the study of controversies online, it could be argued that Mapping Controversies approaches are more controversy-centred and Digital Methods are more device-centred, through the two approaches frequently overlap and individual studies may oscillate between the two poles at different points in the analysis. Mapping Controversies largely instrumentalise media to study controversies, losing the media's role in the process while Digital Methods often, arguably, instrumentalise controversies to learn about media (for a full discussion see Marres and Moats, 2015). But the above examples have suggested that both positions on their own may be prolembatic, at least in terms of 'in action' studies.

Scholars in media studies, indeed, have long argued against the very choice between a focus on media and focus on substantive content: Corner, Richardson and Fenton clearly embrace such a tension in their study of nuclear power as public issue television:

> ‘So the research has a dual 'edge' -- cutting equally into questions about TV as a form of public communication, grounded in particular institutional relations and conventions, and ones about the public meanings which had gathered around nuclear energy at the time. The selection of the energy issue is not, therefore, just an 'example' allowing us to extrapolate off into a general theory of public-issue television, nor is it a substantive focal point in relation to which questions of media discourse are secondary.(Corner et al., 1990: 19)

As I will argue over the following chapters, the attempt to describe the role of media in public science controversies ‘in action’ provides an empirically viable way of keeping these two processes in tension (Marres 2015, Marres and Moats 2015 forthcoming).

What I think this requires is *decentring* devices as the given unit of study. I say decentring because this tension expressed above is asymmetrical. Researchers are dependent on devices to access controversies and have only partial access to them. But this does not mean researchers should lose track of these elusive objects. One of the reasons why I think maintaining this tension is necessary is that we should respect some of the possible scepticism within STS about the role of mainstream media and particularly online platforms in politics and controversies. We cannot assume that media, let alone online platforms, will decisively impact controversies. Although there is no publicisation of controversies which does not also shape them in some way (Hilgartner, 2000), these impacts may be more or less prominent, more or less controversial, depending on the case at hand. This means we should remain open to the possibility that particular empirical phenomena we observe might be emerging from the controversy or, conversely, an effect of the platform. On one hand we do not want to focus on media platforms and routines for their own sake at the expense of the controversy, but at the same time it is important to understand particular effects which may be consequential for the controversy’s settlement. This also allows for the possibility that controversies may not resemble the public science controversies we are accustomed to and start to look like media debates, unexpected forms of politics or protest, in which case they broaden out to issues (Marres, 2015). This may requires expanding our vocabulary for describing what counts in relation to them. Finally, maintaining this tension and granting ourselves the freedom to move between these two poles of the analysis is just as much of a methodological and practical exercise as it is a conceptual one – which I will explore in the next chapter.

04. CONCLUSION

In this chapter, I argued for widening the discussion from online platforms and controversies to take in now classical concerns and approaches in STS and media studies. This diversion was necessary because one possible way of understanding the role of online platforms in controversies is in relation to media: in relation to the circulation of particular representations of the controversy. This was also important because past approaches to media exhibit a tension between studying media content or studying media technologies, or studying media in isolation as opposed to entangled with other media. What these tensions may mask is a more pressing concern for the current study: the tension between studying particular devices and studying controversies with them.

First I argued that ANT-informed controversy analysis and the engaged programme in PUS have historically side-lined the media. It may be the case that this was because media representations are seen as slight or inconsequential in comparison with the meaty material of scientific papers and policy documents. As I will discuss in the next chapter, this is also to some extent a practical problem because media have not been easily traceable in the same way as scientific papers and documents in laboratories or the minutes of meetings in public hearings and other engagement exercises. Participatory platforms may offer the opportunity to study how certain media intervene and shape controversies, but STS does not necessarily have the analytic framework to understand these media on these terms, nor is this its normal remit.

There are some studies within media in relation to 'framing' which points to the work of issue or controversy definition in the mass media, however this work is largely social constructivist and requires further attention to the role of non-humans in stabilising frames as well as the active participation of the controversy in its own articulation. Recent work influenced by STS offers just such a symmetrical focus on the social and technical dimensions of media technologies but does so at the expense of analyses of texts or particular contents of media. It also appears to inherit untenable dichotomies between content and the material and producers and consumers of media

The STS concept of device however allows us to recast the boundaries of the study, the role of audiences and producers and the textual material distinction as empirical topics rather than conceptual problems. But this left us with a different choice between

studies centred on devices, and those centred on controversies or matters of concern. I showed the limitations of both approaches on their own for the present study and proposed that we maintain a tension between device-centred and controversy-centred approaches. We can evaluate whether participatory media shape the controversy in terms of presenting knowledge or definitions and whether or not controversies upset the normal working order of the devices. Finally, though we will be necessarily tied to certain devices, which other media cover the controversy also becomes at stake in the analysis.

This is an approach which adds back in some of the conceptual nuances of ANT into Mapping Controversies but it does *not* do so by limiting the study to particular media or settings. It argues for a certain freedom of movement to trace overflows between platforms. However, and this is the central argument of the thesis, it is challenging to practice this empirical freedom of movement as methodologically our ability to move is both granted by *and* constrained by online platforms. These platforms on one hand allow for the very possibility of tracing information flows between platforms but, as I will discuss in more detail in the following chapters, they also orient the researcher towards certain sorts of analyses, methods and tools as opposed to others. In short they implicitly offer ways of defining the boundaries of the study which become problematic if we maintain this tension between controversies and media. The question for the next chapter is how to utilize these digital traces without being ruled by them, to maintain our empirical commitment to particular controversies which overspill and complicate the articulations put forward by platforms.

III. PLATFORMS AND DIGITAL METHODS: REDESTRIBUTING QUANT AND QUAL

Recent developments in online media are not only said to increase participation in democracy or the media landscape, they are also believed to make possible new modes of research. The internet and particularly, online platforms generate mountains of digital data and may allow researchers to analyse social life at an unprecedented scale and granularity of detail (Manovich, 2012). This inundation of data enables (and might require) the use of large-scale quantitative mapping and analysis techniques to make sense of things, but the techniques and platforms which enable these new methodologies present challenges as well.

In the previous chapter I argued that in order to study the role of online platforms in controversies, researchers should maintain a tension between studying controversies and studying devices: leaving open the question of which devices are most central to the controversies, and also which sorts of media effects and processes are most relevant to its development. In other words, I endorsed the argument for a certain freedom of movement, which has long been insisted on in the ANT-strain of controversy analysis. Yet this freedom of movement is both further enabled and complicated by the rise of online platforms. While they make available more formatted data than the Web in and of itself, they also, arguably, exert more influence on online research: platforms may over-determine the boundaries of the study; they are infused with metrics and representations which may incline us to accept their definition of what is important; finally they may, as I described in the last chapter, enforce the analytical separation of content and materiality: what travels and the infrastructure and arrangements through which it travels (Star, 1999).

In this chapter I will evaluate and discuss both qualitative and quantitative techniques for researching controversies with online platforms. I want to stress that what I am proposing is not a balancing act between quantitative and qualitative or another call for mixed methods. I will not phrase my contribution in terms of the extensive debates about 'quant' and 'qual' (Fielding and Fielding, 2008; Hammersley, 1992)

or possible bridges between the two (Glaser and Strauss, 1967). STS thinkers, have a unique take on methods, drawing on their understanding of scientific practice, which transcend these classic divisions. Also, in the empirical chapters that follow, the lines separating automated tools and manual analysis will become blurred in practice. However, many debates about online platforms or big data currently fall along inherited 'quant' and 'qual' lines, which obscures some of the more subtle tensions at work. In the first section I will start with a discussion of debates around digital data and discuss the claim that the new data sources offer an occasion to rethink inherited and entrenched divisions between quantitative and qualitative methods and micro and macro scales of analysis (Venturini and Latour, 2010).

Next, I will show how such a proposal would work with web data – static webpages and hyperlinks – and how virtual ethnography and hyperlink analysis allowed for both the charting of controversies and studying devices. But the emergence of larger online platforms, as opposed to discreet websites, I argue, complicates this work. Digital Methods techniques offer some help in this area because they make the formatting work of devices central to the analysis but in the same way as in other STS accounts of the role of media in controversies, there is an on-going tension between instrumentalising platforms to study controversies and using controversies to study platforms. Also Digital Methods does not clearly articulate a program of textual analysis or observation, so I will offer some techniques borrowed from past work in STS to better flesh out the 'in action' approach.

Although this survey of existing approaches is necessary to ground the empirical studies, the central argument of this chapter is that if the object of study is decentred by controversies, in addition to Digital Methods and device-centred techniques, it may also be necessary to decouple our methods from these devices, partly with a renewed focus on qualitative techniques and secondly with different types of data visualisations which emerge from some of the less obvious and accessible devices and technologies which constitute these platforms. What I mean by this will be progressively elaborated over the course of the empirical chapters but for now the relatively simple point I want to make, using the work of Gabrielle Tarde, is that it is important to distinguish between what is important to analyse from what is easy to analyse.

01. BIG DATA AND QUALI-QUANTITATIVE METHODS

Despite the association of ANT and PUS with ethnographic and also semiotic techniques, which were used to interrogate the closed world of science, researchers in STS have, from the very beginning, experimented with quantitative techniques and visualisations: from simple graphs of scientific citations over time in *Laboratory Life* (Latour and Woolgar, 1979) to the development of co-word analysis in *Mapping the Dynamics of Science and Technology* (Callon et al., 1986 discussed later), these approaches were an extension of the qualitative work of analysing texts and observing laboratory practice, taking in a longer historical vantage and a more macro scale of analysis. This was possible because science, unlike other areas of social life, is relatively well documented and structured in systematic ways: particularly in the case of scientific articles. The rise of internet data has, for several thinkers, finally granted the possibility that these techniques, emerging out of scientometrics as well as controversy analysis, can be applied in other domains.

The internet presents sociologists with staggering amounts of 'transactional data' (Savage and Burrows, 2007) or 'digital traces' (Latour et al., 2012) such as hyperlinks, timestamps, log files, comments and texts which can be studied through both quantitative and qualitative techniques. Although there have been debates about internet research methods from the very beginning, current discussions of methods for researching online data have become wrapped up in what is called 'big data' (Kitchin, 2014).[61] In sociology, big data is often conflated with 'social media' data, the most readily available source, but the term has a wider usage, which includes open government data and largely proprietary data accumulated by private companies.[62]

---

[61] Big data originated as a computing term (Manovich, 2012), which referred to data which either required multiple servers to process or particular software packages like 'R'. It has since been taken up as a market research term by corporations and governments interested in using data to solve problems and make evidence based decisions. As Kitchen notes, defining big data in terms of its size is not particularly helpful; what distinguishes big data from what has been retroactively dubbed 'small data', is the famous three V's: 'Volume' (amount), 'Variety' (heterogeneity) and 'Velocity' (made available in close to real time) to which Kitchen adds 'Exhaustivity' (the aspiration or illusion of comprehensiveness), 'Resolution' (fine grained detail), 'Relationality' (easily combined with other data sets) and Flexibility (open to multiple uses).

[62] Another way to think of this data is in relation to social science research. Big data or more broadly 'transactional' data or 'by-product' data is not collected actively by researchers as in the case of surveys, questionnaires, focus groups etc, but amassed *incidentally* as various interactions take place (loyalty card transactions, purchases, doctors visits, public transport, comments). This is not to say that this data collection is completely non-intrusive or without social research goals, but that data collection and analysis is being led by the private sector instead of the academy (Savage and Burrows, 2007). Marres (2012b) qualifies, however, that there has always been traffic between the social sciences and the private sector, from the Mass

Regardless of the extent to which big data can be seen as fundamentally new, it has created a very real uproar in the world of research. The term has presided over a push toward the use of computationally advanced techniques both in marketing and social sciences and digital humanities. These include computational social science (Lazer et al., 2009), the building of social science models borrowed from natural science which also includes machine learning (Baysean) algorithms which purport to detect patterns in data, either with or without the aid of a human coder, and which 'improve' in their functioning the more data is fed to them.

This has led several thinkers to launch critiques of big data approaches: that they present problems for access and ethics (boyd and Crawford, 2012) or court researchers into asking reductive and easily answerable questions (Uprichard, 2013; Vis, 2013a). Traditional quantitative researchers also have posed important questions about the sampling bias of given data sets (Sloan et al., 2013) and validity (Tufekci, 2013).[63] However, while it is healthy to be sceptical of the claims of brute computing power or complex maths to easily describe social life, without theory or contextual understandings to back it up, this does not mean one should abandon computational techniques. Given this 'data deluge' (Kitchin, 2014), qualitative researchers may now *require* some sorts of quantitative techniques to locate informants and define the boundaries of their study – otherwise they will be needlessly confined to those corners of the web they can manually read and comprehend.

As I will explain, STS scholars are also critical about the provenance of this data though they are not afraid to engage creatively and reflexively with it: STS scholars tend to see potential biases (Driscoll and Walker, 2014) of media or platforms as something to be studied rather than as a problem to be corrected or adjusted for. It should also be said that their use of automated tools is different from traditional quantitative approaches for which the end result is numbers or models. They use them in a visual, interpretive

---

Observation project to focus groups and opinion polls. What is actually significant about big data is that it redistributes the capacities of those involved in the research: potentially giving partial control of the interpretation of data to technologies or even the research subjects themselves who may define what is relevant to the study in new ways.

[63] While concerns about social media bias are very real, some of the language of validity and sampling is more relevant to research which seeks to use digital data to make statistical claims beyond the data: for example about the demographic characteristics of Twitter users, or deploy samples of Twitter data to make claims about a population.

way in the service of creating *descriptions* as opposed to demonstrating causality (Latour, 2005).

Quali-Quantitative Methods

So big data has largely exacerbated tensions between quantitative and qualitative researchers, but others have argued that the growth of, particularly, internet data, may actually offer an occasion to question the division of labour between the two (Hine, 2005; Sack, 2000; Venturini and Latour, 2010). Latour and Venturini argue that, statistics and ethnography for example, produce different ontologies of 'micro' and 'macro' which are fictions, effects of these divergent methods. However, in two related papers they propose that the digital finally allows for reading data in a 'flat ontology' in the style of ANT, allowing the researcher to zoom seamlessly from the individual out to the aggregate.[64]

Latour, Jensen and Venturini (2012) specifically offer the example of linked social networking 'profiles', which are composed of links pointing to heterogeneous institutions, actors and objects. The authors are here drawing on the alternative sociology of Gabriel Tarde, who has been adopted as a predecessor of ANT (Latour, 2010). In contrast to Durkheim who believed that society existed *sui generis,* independently of individuals, Tarde saw the social composed of relational monads (following Leibniz) which are defined only through their relation to each other. Significantly for adherents of ANT, these monads can be human or non human. Although I will not elaborate on Tarde's work here, at the end of this chapter I will use a somewhat different reading of him to point the way to a different approach to digital data.

In Latour's example, an academic is defined by qualifications from an institution but the institution is defined by the sum of associations with former graduates such as him: the associations are bi-directional. The institution does not contain the professor or exist on another scale, it is on the same scale, another monad, which is also a quality internalised by a host of individuals. These monads and their associations can be

[64] Because they are unable to fully achieve a flat ontology, they refer to their solution as occupying a mezzanine level between the two existing scales (1.5 level). Drawing on an earlier project (Latour et al., 1992) this could also be described as moving between previously fragmented qualitative 'micro-theories' like creating bridges between an archipelago of small-scale case studies.

scrutinised individually or, using network analysis software, mapped to visualise wider relationships and groupings. Elsewhere, but also in relation to digital data, they refer to these combinations of methods (networks in conjunction with qualitative analysis) as 'quali-quantitative' methods (Venturini and Latour, 2010).

This is an important proposition and one which is particularly pertinent to the study of public science controversies online, where the object so easily spans what we think of as the macro and micro scales of analysis: it is particularly hard to situate individual posts, messages and utterances within rather more tenuous constellations of actors and larger controversy dynamics. It should be understood that Latour and Venturini are using this example of an academic profile to make a wider point about social research rather than to outline a robust methodology for studying the internet, but I will use this as a helpful starting point to pose some questions about quali-quantitative methods and how they might relate to the study of controversies or issues.

It helps to think of this proposal through what Law, Ruppert and Savage describe as the 'double social life' of methods (Law et al., 2011). Methods are social, in one sense, because they have histories: they exist and are made possible through assemblages of networks of institutions, research subjects and other methods, which must be taken into account. So it should be noted that the choice of network analysis in the above example is not an innocent move. Forms of network analysis have a particular currency both in STS and in the field of web studies and social media research, something which we may wish to question. These historic pairings between network analysis and particular devices like profiles are not given (Marres and Gerlitz, 2015) yet they are also not arbitrary. This is because methods will always be *entangled* with various devices being studied, which have certain socio-technical capacities for research. In other words, network analysis in this case, becomes feasible due to the particular affordances of online profiles as a device.

This entanglement is not specific to web data or platforms, but is particularly important to acknowledge in these cases. This is because to study new platforms normally requires querying or 'scraping' the platform. Scraping is the process of obtaining structured data either manually or through a platform's API (Automated Programming Interface)[65] There are numerous cautions about the use of scraping and, in particular, APIs (boyd and Crawford, 2012; Manovich, 2012; Savage and Burrows,

[65] For a full discussion see (Marres and Weltevrede, 2013).

2007). The main concern is that the structure of data and the access granted by APIs may incline researchers to certain sorts of questions and methods as opposed to others. As Marres and Weltrevrede note, this dependency may prompt sociologists to pursue 'real-time' or 'trending' data (Marres and Weltevrede, 2013). So it is important to understand the politics of particular devices that our methods are enmeshed with not only as an interesting phenomenon in its own right, but also because they may slant our research. Because these entanglements are inevitable, one tactic of STS researchers has been to *align* their methods as much as possible with the devices under examination, and in this way, make the formatting work of devices part of the investigation.

The second way in which methods are 'social' is that they *produce* the social. Just as inscription devices in the natural sciences produce the phenomena they purport to describe (nature), social science methods, like other devices, produce or rather *perform* entities like the nation, or forms of reflexive subjectivities (Savage 2010) and can be interrogated along these lines. This point requires some unpacking. As discussed in the last chapter, the performative effects of devices depend on how pervasive or successful they are: a network diagram in an STS journal is not going to move the world in the same way as a public opinion poll featured prominently on the nine o'clock news (Osborne and Rose, 1999). But this point becomes relevant to the extent that methods like network analysis are successfully aligned to particular devices like profiles. These devices *format* social life in a particular way: they create opportunities for associating, which were not there before. Digital Methods approaches have indeed argued that devices such as profiles or other forms of networked content are themselves methods (Rogers, 2013b): they are ways of knowing and navigating social life and using network analysis in social research is a way of 'repurposing' these methods. This could be described as a performative approach to online data analysis in which the interventions of the researcher is rather minimal, they 'follow the medium' in a parallel sense to the slogan 'follow the actors'. This is a second reason offered for why it is appropriate to align methods with devices, because we do not, through our methods want to impose incongruous pre-conceptions of the social or politics on them, but rather study the particular forms of sociality they enact.[66]

[66] Rogers (2012) illustrates this point with a twin history of shifting technologies (hyperlinks, lists, directories, search) and social science conceptions of politics they produce (the round-table, the network, the sphere etc.).

Now for the time being, I am going to gloss over the issue of alignment between the device of online profiles and the visualisation of this data by social scientists as a network diagram (Marres and Gerlitz, 2015) but later in the chapter, this relation between method and object will become important to consider. But for now, the point I want to raise is that network analysis, both in the sense of network visualisations and the Tarde-influenced tracing of associations, may align well with networked profiles but may encounter problems when online platforms are involved.

The second point I want to flag is that Latour and his co-authors mostly instrumentalise or repurpose these hypothetical online profiles to map associations between actors.[67] This based on the premise that technical artefacts like hyperlinks are traces of the movement of actors like a kind of infrastructure.[68] Yet the very same data and networks could be used to explore the device of profiles itself and some of the *politics* of linking. Profile connections, just like hyperlinks or scientific citations, create forms of association in which users and institutions must maximise their visibility or authority at the expense of their competitors. In the last chapter I argued that we needed to attend to both – what the devices disclose about controversies and particular forms of association, enacted by the devices – and that one way of doing so is to study controversies 'in action'.

To summarize this admittedly complex section, I made the point that online data, according to certain thinkers in STS, may allow us to finally resolve divisions between 'quant' and 'qual', macro and micro even though these divisions are hardwired within the social sciences. I explained the general STS position on methods: that they are embedded in larger 'method assemblages' (Law, 2004) which include the objects and devices and research subjects being studied, and they are performative in that they enact certain forms of social life. STS scholars then tend to study methods in the world or latch on to these methods rather than bring their own to the table.

However, I proposed a few cautions: while network analysis is an appropriate way to study profiles and, as I will discuss, web data, it is worth asking if it is the most

---

[67] Although networks should be a great fit for an actor-network theory approach, Latour had previously (2005) said that the network being traced as part of an ANT ethnography would never really look like a network, it could not be drawn as such,

[68] Or to use the famous phrase 'Technology is society made durable' (Latour, 1991): frequently crossed paths will become paved, lost hotel door keys gain cumbersome fobs to prevent them being taken away. These stubborn artefacts which, are costly to unmake, become a material instantiation of associations which came before.

appropriate method for studying platforms. This also raises the question of what sort of qualitative techniques are to be paired to platforms. Interestingly, the qualitative side of quali-quantitative methods is not really defined: is this an ANT informed ethnographic technique or discourse analysis? Finally, following the tension identified in the last chapter, can these methods be used to both examine the contents of devices as well as turn back and examine the devices themselves as socio-technical assemblages?

In the next section, I will discuss some existing STS-informed approaches to controversies on the web: Virtual Methods and Issue Crawler in which such a quali-quantitative programme was made feasible and then turn to new approaches to platforms. I will also use this discussion to raise the three interrelated problems identified in the beginning which have haunted this thesis so far: the problem of indeterminate boundaries, the influence of relevance defining metrics (which have implications for the balance of 'quant' and 'qual') and the tension between content and materiality.

02. VIRTUAL METHODS AND HYPERLINK ANALYSIS

While some of the earliest studies of the web concerned relatively cohesive 'virtual communities' (Rheingold, 1993) focused around topics of interest on Usenet forums or MUDs (boyd, 2009) the development of the world wide web and user-created web pages fragmented and distributed this stable object. *Virtual Ethnography*, proposed by Christine Hine (2000) was introduced as a reflexive answer to this problem.[69] This became part of what is known as the Virtual Methods approach (Hine, 2005) which adapts existing social science methods (survey, interview, social network analysis) to the web, capitalising on opportunities while minimising constraints. Virtual ethnography problematized the traditional object of 'communities' and the extent to which individual websites can ever be the appropriate delimiter of a study (Guimaraes, 2005; Hine, 2005).[70] Hine instead recommends a focus on 'topics': in her study, she

[69] I will focus on Virtual Ethnography in this chapter as opposed to Digital Ethnography (see Murthy, 2008) which normally explores the relation between online and offline work, facilitated by digital media including offline analysis packages such as NUD*IST and Nvivo. Virtual Ethnography is perhaps more questioning of the status of digital data and it also has been applied to the study of public controversies or 'media events'

[70] Virtual ethnography approximates an ethnographic approach by observing live interactions, studying webpages as both a 'cultural artefact' and as a 'space of interaction'. Although this does

investigates a media event regarding the trial of Louise Woodward and studies the proliferation of websites and internet forums which are devoted to covering the trial.[71] Topics in this usage do not entail the constitutive uncertainty of issues or the knowledge content of controversies but they do define the object as separate from but mediated by online media.

Yet whether defining the object as topics, issues or controversies *locating* the object is a central problem which is contingent on the very devices being studied. For example Hine starts her analysis with a search engine to locate relevant pages through search terms. She acknowledges the potential role of actor's search engine optimisation (strategically attempting to improve one's rankings) in possibly skewing her ethnography – that she may not see the less professional or Google-friendly sites. If a study starts with search engines or alternatively social media sites to gather starting points, then it becomes shaped by how the device, including technologies, algorithms and the input of users of the device formats the controversy and defines what is relevant or popular.[72]

Understandably, the second section of Hine's edited volume *Virtual Methods* (2005) is devoted to the problem of locating studies. Dodge's (2005) contribution for example proposes that various quantitative 'mapping' techniques, which can include a number of ways of representing information visually, can offer strategies for qualitative researchers to grasp larger patterns, spot holes in the data and make sense of formations not visible through individual postings or pages. One of the most common mapping techniques for assisting ethnography has been hyperlink analysis. Beaulieu (2005) suggests that hyperlink analysis allows ethnography to scale up to extend the process of tracing links which it performs already. The two methods actually parallel each other. This can take one of two forms: either data is gathered ethnographically by

---

not exactly map on to my distinction between studying platforms for their content and studying platforms as an object in their own right, Hine's point is that the ethnographer must be prepared to switch modes of analysis (between analysing texts or observing practices) depending on the situation.

[71] However, Hine also places the emphasis on individual users as the focus of virtual ethnography, which may make sense in the case of webmasters maintaining static webpages, but less so in social media platforms where particular users may only have a passing interest in the topic.

[72] To the extent that this popularity is implicated in power asymmetries between participants or particular media frames rising to the top, this is important to study, however the most pertinent data for the controversy may not be the most popular.

tracing links and later aggregated into a network or a program such as a crawler is used to trace networks of links automatically to guide further ethnographic tracings.[73]

The former might be exemplified by the e-Diasporas project (Diminescu et al., 2011). Experts on a particular diasporic group use Navicrawler, a Firefox add-on, which records pages visited and their out-links. The researcher collects the network data in the normal process of tracing links and the macro results can then be visualised and analysed as a network graph.[74] Arguing for the latter approach, Howard (2002) proposes that social network analysis can be deployed for locating 'field sites' from which to launch ethnographic studies, in a similar way to other offline, ethnographic tactics for 'constructing the field' (Amit, 2000). Park and Thelwall (2005) use automated crawling and social network analysis as a way of locating the key actors or informants with the aid of statistical measures such as 'degree centrality' and 'betweenness', which are ways of identifying nodes with the most connections or strategically important positions (see Hogan et al., 2008 for an explanation of network analysis metrics).[75]

In the case of sprawling controversies or matters of concern, which present a challenge to ethnographic tracing, it makes sense to let the mapping guide the corpus selection, though in practice this will always be to some extent an iterative process in which qualitative studies inform quantitative mapping and vice versa. But as argued earlier, it is also important to interrogate the status of hyperlinks. As many virtual ethnographers have discovered, links can be interpreted as inferring authority, providing functional connectivity or fulfilling more symbolic functions depending on the context (Beaulieu, 2005). Their use can also vary wildly depending on how they are qualified on the page (positive, negative, ambivalent etc.)(Rogers, 2012) This is why, when instrumentalising hyperlinks as a way of locating 'relevant' or 'authoritative' actors it also seems crucial to question the relevance defining role of hyperlinks in

[73] Crawlers are software tools which gather and map hyperlink patters by scraping websites for URLs (see Bruns, 2007 for a full discussion)

[74] This is perhaps the closest approximation of the approach proposed by Latour and his co-authors, but as Diminescue acknowledges, the project and the Navi-crawler tool are 'Web 1.0 focused' — they work for personal websites and blogs but researchers must perform different analyses to understand diasporic groups' use of social media like Facebook and Twitter.

[75] But using mapping techniques does not require putting so much weight on the metrics of network analysis, which to some extent are formalist: measures of centrality depend on *structural* positions in the network which assumes that each link is of equal value. This imports assumptions about links as markers of authority in the sense of scientific citations or points of infrastructure for communication in classic social network analysis which may not always apply to hyperlinks.

formatting controversies. This becomes increasingly important as the web came to be defined by search engines like Google which incorporate hyperlinks, as markers of authority, into their algorithmic ranking of sites, leading users to game the system (Rogers, 2004).

Issue Crawler

One approach which more explicitly explores the politics of hyperlinking is the Issue Crawler (Marres and Rogers, 2000, 2005). While other forms of hyperlink analysis are geared to social networks or communities, Issue Crawler's object is a particular issue. The tool starts from a small corpus of websites explicitly devoted to the issue, usually compiled by an expert, and crawls the out-links of those pages and their out-links for a set number of iterations. The larger corpus is then reduced to sites which interlink using the measure of 'co-link', based on the scientometric measure of 'co-citaton'. The result is a network of sites called an 'issue network'. This could be seen as critique and revision of deliberative models of politics focused predominantly around discourse or information sharing. It also however is a critique of social understandings of political communities – when these heterogeneous actors may have antagonistic or ambivalent relationships, defined only by shared commitment to an issue (Marres, 2006).

But today these already elusive gatherings may be even harder to detect. Issue Crawler was developed at a time when most sites: NGOs, corporate, government, activist (Marres and Rogers, 2005) or blogs (Bruns, 2007) were all roughly symmetrical in function and layout – 'static web' pages maintained and updated by a webmaster. Crucially they all had some form of links page or 'blog-roll', which referred the visitor to friendly (or sometimes antagonistic) sites.[76] But today researchers using Issue Crawler are frequently confronted with new online platforms.[77] For example, when I performed a crawl around the issue of nuclear power in the UK, the resulting network did not only include individual actors, organisations and institutions selectively linking, but online news websites and social media (Figure 2).[78]

---

[76] Digital Methods researchers have themselves asked the question Is Issue Crawler Web 1.0? Available from: http://blog.digitalmethods.net/2010/how-web-1-0-is-the-issuecrawler/ (Accessed 11 September 2015)

[77] It has long been recognized that Issue Crawler studies will encounter online news nodes in issue networks (Marres, 2002).

[78] This issue-network (Figure 2) was produced starting with an expert generated list of relevant activist organisations government branches and energy companies. These starting links were run through the Issue Crawler tool on the 'co-link' setting for two iterations – the crawler

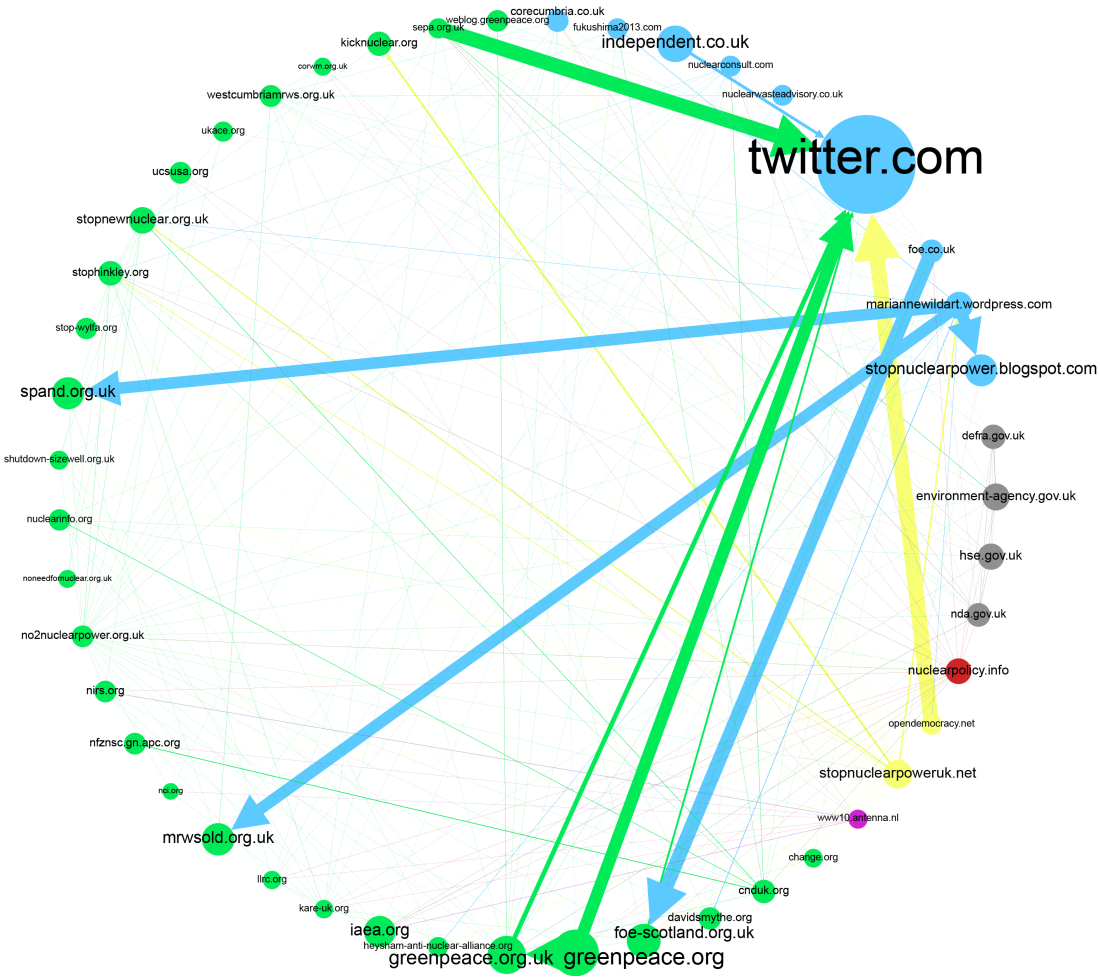


Figure 2. Issue Network https://goo.gl/FvRqIx: Issue Crawler map of the nuclear power issue in May 2013. Note the prominence of Twitter and broadsheet newspapers in the network.

Figure 2 contains the expected mix of NGO, government and corporate actors but also, video sharing site Vimeo, *the Independent* newspaper and especially Twitter stand out. Because Issue Crawler is designed to aggregate host URLs as opposed to individual articles or pages or users, these media and social media sites become 'black boxed'. The map cannot disclose which particular Twitter users or hashtags relate to the issue. Now in other forms of hyperlink analysis, Twitter might be understood as a highly 'relevant' or authoritative actor, yet such a map can also prompt the researcher to

---

scrapes the starting pages for links, and scrapes the links on the resulting pages. Issue Crawler then reduced the larger corpus of links to those sites that are also linked to by the network. The results are then imported into network analysis tool Gephi where they are spatialised as a round-table, colour coded by domain type (.com .org .uk) and resized according to 'authority' (the number of links from the network).

investigate the hyperlinking practices, which enable this finding. For example, Twitter does not receive in-links from most of the network but *many* links from a few sources – because these sources mostly have a Twitter button or Twitter widget on the side bar of *multiple* pages. This demonstrates how different sites use hyperlinks differently. But the very appearance of Twitter, social media and news sites in the network also raises questions about the integration of Twitter and News sites into the practice of certain activist and corporate websites. It even suggests that they may be increasingly geared to 'newsy', real time events. Issue Crawler maps could just as easily have quantitative metrics applied to them but they can also be analysed in a way which raises questions about the technology of hyperlinking and guides further research.

Together, virtual ethnography and forms of hyperlink analysis can form a viable quali-quantitative approach which has great purchase in investigating controversies in certain aspects of the web. They are complimentary and iterative and both the quantitative and qualitative aspects can be used to both study an issue or topic *and* study the web-specific organisation of that topic. Yet both methods may be complicated by the rise of platforms.[79] Now, I want to be careful to not phrase this problem in terms of periodising the web. As Rogers explains (2012), logics which organise the web do not supplant each other but co-exist and even interfere with each other (Rogers, 2012). Anne Helmond's recent dissertation examines the 'platformisation' of the web, how platform specific objects become increasingly integrated into the fabric of the web as part of an 'ecology' (Helmond, 2015). So while hyperlinks were once the central currency of the web (Rogers, 2004) and search engines rankings like those of Google were built on this logic, in addition, now platforms like Youtube, Twitter, Instagram and Facebook have different platform-specific devices such as 'likes', 'hashtags', pages and metrics (based on keyword rankings and friendship or follower networks) for defining the relevance of content and distributing it to particular users. So instead of a multitude of individual sites, much interaction is fed through and managed by these 'mega sites' which may warrant separate studies and methods and thus have the effect of fragmenting the object of study.[80]

---

[79] Both approaches presume to some extent that hyperlinks are the primary technology which links disparate aspects of the web. But it is indeed possible, as Elmer argues, that the hyperlink is no longer the primary organising technology of the web (Elmer, 2006). What may be necessary then is an alternate pairing of quantitative mapping and qualitative analysis which does not rely so much on hyperlinks.

[80] Although I do not want to associate platforms too strongly with either 'Web 2.0' it is worth raising two points about the status of domains and hyperlinks. O'Reilly (2005) makes the observation that, while Web 1.0 or 'static' pages were manually updated by a webmaster, 'Web

03. DIGITAL METHODS

One way of attending to this proliferation of platforms and their different logics for organising and structuring data is to focus on the relevance defining of the platforms themselves. Digital Methods (Rogers, 2009 and 2013) introduced in the last chapter is a research programme which claims to 'repurpose' the 'dominant devices' of the web, to study social life.[81] Researchers at the Digital Methods Initiative (DMI) in Amsterdam have developed an arsenal of *platform-specific* tools, available in browser-based interfaces which 'follow the medium' or the 'methods' built into the medium.[82]

Digital Methods often use 'platform', 'device' and 'medium' interchangeably because their approach sits between media studies and STS. Digital Methods are not completely dissimilar from computationally advanced or big data approaches, but there are two key differences. First, they employ relatively simple techniques (word frequency, co-occurrence) and simple visualisations (word clouds, bar chats, pie chats, network visualisations) as opposed to advanced statistical measures such as regression analysis. This allows researchers to easily trace results back to the devices and issues being studied. Second, while Digital Methods tools normally start with an expert-generated list of sites, actors or search terms, DMI researchers never curtail or alter the data after the fact, removing 'outliers' or accidents of the process, because these artefacts may become findings themselves (Rogers, 2013b).[83] One of the goals of Digital Methods is thus studying the 'bias' of the medium, rather than taking bias as something to be weeded out.

2.0' pages are *dynamically* updated by scripts, automatically generating content. Another technical feature is the rise of the 'permalink' (unique URLs which reference individual posts or articles) as opposed to 'Web 1.0' pages in which new posts replace old ones or move to the top of a scrolling front page on older blogs. What this suggests is that individual posts or articles may be a more appropriate unit of analysis than host domains (the individual post or article rather than the web site).

[81] Digital Methods places itself in opposition to Virtual Methods, like Virtual Ethnography which modify existing social science techniques. But since many internet platforms are inspired by social science methods (Google's ranking algorithm is inspired by scientometrics) Digital Methods of course still deploy traditional social science methods like social network analysis. The difference is more about the closeness of fit between the data and the method used.

[82] Available here: https://www.digitalmethods.net/Dmi/ToolDatabase (Accessed 29 Sept 2015)

[83] This could be justified in terms of 'following the actors' – allowing the definition of the object of study to emerge out of competing definitions of research subjects rather than the researcher's assumed definitions.

It is worth contrasting this approach to data with similar tools developed at the Média Lab at Sciences-Po as part of the Mapping Controversies Project. Hyphe (http://www.medialab.sciences-po.fr/fr/tools/hyphe/), one of the latest tools, which builds web corpuses in a similar way to Issue Crawler and ANTA (Venturini and Guido, 2012 discussed later), which analyses text, both encourage the researcher to edit the resulting maps based on researcher-defined criteria and thresholds, while Digital Methods would leave this relevance-defining to the device.

For example, one of the key DMI tools is the Google Scraper, which quantifies the appearance of certain key terms on a given list of web sites.[84] This is useful for studying differences between country-specific Google engines (google.co.uk, google.fr) (Rogers et al., 2013) or search engine demarcated web-spheres such as the (Google Blog Search) 'blogosphere' or (Google News)'news sphere'. Google can be queried to produce a list of websites relevant to an issue and these websites can then be queried with Google for key terms. The results are then visualised as word clouds for each website (or grouping of sites) sized by the frequency of occurrence of words.

This is like an automated version of quantitative media content analysis (Philo, 2007) in which the comparative frequency of words can highlight 'partisanship' of media. But Digital Methods also exhibits a, not unproductive, tension between studying controversies (or issues) and studying platforms. Digital Methods is mainly focused on capturing broader societal trends but the analysis is as much about the device as the issue: in this case how Google idiosyncratically ranks with its algorithm. In other words there is a 'technicity of content' (Niederer and van Dijck, 2010) because a platform's architecture, technologies, non-human 'bots' and scripts all participate in producing articles, posts etc.

Because the data is not 'edited' there are accidents of the process, search terms may generate pages related to *different* issues (nuclear power or nuclear weapons?) or artefacts of dynamically updated webpages which reveal how Google's algorithm or website crawling spiders work. Instead of merely instrumentalising Google to locate

[84] Google is one particularly useful device to study because it allows access to other platforms. Rogers explains (2012) that the influence has been so pervasive that Google and other search engines create 'engine-demarcated spheres' such as the 'news sphere' or the 'blog sphere'.

relevant websites, it highlights the process by which Google deems websites 'relevant' or 'popular' with respect to the controversy at hand.

These researchers have increasingly turned their attention to platforms (Rogers, 2013b) by repurposing particular objects or devices featured in them. Wikipedia has been analysed using In:links between articles and edit statistics (Currie, 2012) or images (Rogers et al., 2013); Facebook pages have been repurposed to create like networks (Rieder, 2013), while Twitter analysis has focused on, among other objects the Hashtag (Borra and Rieder, 2014; Marres and Weltevrede, 2013).

Digital Methods represent a reflexive way of dealing with platforms, but I want to raise a few interlinked points about these techniques, given that in the last chapter I argued for decentring the object of the study from devices toward particular controversies. As I suggested, the object of controversies may pose questions about, at the very least, which devices to repurpose.

Digital Methods are by definition confined to particular devices, which is what gives them their analytical strength, but controversies always overspill the particular venues in which they are debated. The language used on Wikipedia may be mostly driven by news or the activity in news and blogs may be following revelations on Twitter. There will also be multiple platform-specific objects speaking to the controversy within each platform (hashtag or page or article). As many Digital Methods researchers have pointed out, which identifiers of the controversy are most relevant, is *itself* part of the controversy: on Wikipedia, Gerlitz and Stevenson (Gerlitz and Stevenson, 2009) see in particular the process of forking or splitting articles as a Wikipedia-specific way of diffusing controversy or strategically repositioning it (see also Tkacz, 2010 on the politics of 'forking' articles). This might necessitate iteratively shifting the analysis from the given starting points.

This means that researchers can, and should use multiple Digital Methods approaches, based on different devices, to monitor controversies, but this still makes tracing information flows between them difficult. This could be aided by the free-form tracing of Virtual Ethnography but, and this is my second point, Digital Methods does not articulate clearly a qualitative component, though they certainly advocate close reading and 'spending time with your data' (Rogers, 2013b). Finally, the focus on platform objects, ranking them and mapping them is important to the extent that these devices

are treated as a topic to be studied, there is also a danger, expressed earlier, that in following them we adopt the device's frame of reference as the only definition of what is significant about the study. I will deal with these points in reverse order.

Relational Measures: From 'Liveness' to 'Liveliness'.

Firstly, I will address the problem of metrics over-determining the study.
Imagine that the researcher starts the study using Digital Methods inspired techniques to monitor multiple platforms (including perhaps Google and various participatory platforms: Wikipedia, Youtube, Reddit, Facebook, Twitter) based on certain delimiters or objects (key words, pages, hashtags etc) each of which have different metrics which denote frequency of activity: number of edits, articles, mentions etc. Spikes in activity indicate a sort of gathering of actors in particular platforms or spaces at particular times which might warrant further investigation. These are very blunt measures which may, roughly, indicate controversiality. However frequency and activity measures are not enough because activity measures can either be an *intensification* of the controversy or an *effect* of media-specific activities related to boosting readership or hype (Marres and Moats, 2015; Marres and Weltevrede, 2013). Some of these activities will be strategically important for promoting positions within a controversy but others will be more relevant to studies interested in media themselves.

Marres and Weltrevrede refer to frequency measures, which often map on to the 'real time' or trend based popularity measures of social media as measures of 'liveness' and they advocate instead ways of looking at 'liveliness' or the relational dynamics of shifting content. One way to accomplish this is through co-word analysis, based on the work of Callon et al in *Mapping the Dynamics of Science and Technology* (Callon et al., 1986). The original purpose of this technique was to map changes in scientific fields through the visualisation of large corpuses of abstracts. Key terms which stood for established and emerging fields of research were seen to form around shared 'problems' or objects which they shared. Co-word works by analysing a corpus of texts for the co-occurrence of terms. Co-occurrence can be defined as taking place within a sentence, a paragraph, a document or a pre-defined distance of words (i.e.- appear within 5 words of each other). When words co-occur they are seen as associated and the more times they co-occur the stronger the association. The results are visualised as a network and can be clustered according to groupings of words using various

procedures (Chateauraynaud, 2009; for more information see Danowski, 2009; Hellsten et al., 2010)

Importantly, co-word moves past frequency by revealing more than just which words are popular or 'trending' on a specific device, co-word is *relational* and reveals dynamics between words (Marres and Weltevrede, 2013).[85] In Callon's version however, this insight only becomes relevant in networks over time, to determine which words in certain time slices are stable and which are changing and from which networks of words they emerge.[86] What is also interesting about such relational methods (including other forms of network analysis) is that it becomes easier to tease apart effects arising from a controversy and those arising from media, though the two can never be fully disentangled. On Twitter for example, DMI researchers have developed the technique of co-hashtag analysis (Marres and Gerlitz, 2015) which visualises the relations between hashtags which occur in the same tweets. Most of the hastags displayed in a network will refer to the content: the event or controversy being covered (Marres and Gerlitz, 2015), depending of course on which key words are used to define the corpus – #nonukes, #EDF, #Hinkley, etc. are all hashtags related to UK nuclear disputes. Yet other hashtags like #FF may also be linked to these. #FF stands for Follow Friday which is a campaign that suggests Twitter users should follow new users on Friday – it is a hashtag specific to Twitter. While this could often be dismissed as part of Twitter's self-promotional culture and thus less significant for the analysis, some of these platform-specific promotional strategies might result in the promotion of particular articulations of nuclear power over others. As long as the researcher does not clean the data with a preconceived idea of what is significant, these accidents of the process can reveal the formatting of the device.

---

[85] Hellsten, Dawson, and Leywendorf ,who are also analysing scientific controversies in the media, see co-word in relation to media framing, discussed in the previous chapter (2010). They argue that pure word frequency in the style of media content analysis can capture 'explicit framing' but what co-word also captures is what they term 'implicit frames'. I am sceptical about the potentially formalist assumptions about language structure this claim is based on and the idealised use of media frames but the point stands that relational measures reveal patterns which are lost in frequency based measures.

[86] But one key difference between mapping scientific abstracts and internet texts is the extreme variation in length, style and presentation. A text means something very different in different platforms: comments, tweets and blog posts are all of different lengths and this can skew the visualisation of co-word maps. In a Digital Methods context, (Marres and Weltevrede, 2013; Rieder, 2012) have applied co-word to Twitter where the definition of co-occurrence is within a Tweet. This is convenient because Tweets are roughly the same length (140 characters or less). (Danowski, 2009) in contrast explains how to apply co-word to a variety of media types using a stable distances of words rather than a document as the delimiter of co-occurrence. But this means loosing any differentiation in a corpus of grouped texts.

Another way to highlight this format work would be to include, not just textual data, but other sorts of objects in these maps. Venturini and his co-authors (Venturini et al., 2014), this time not in relation to digital media, use a co-word inspired tool called ANTA (Venturini and Guido, 2012) which uses *bi-partite* networks: containing two types of nodes. Using a data set of the recorded minutes of UN climate hearings, they visualise relationships between both key terms and the documents they appear in. Key terms are not connected to each other but connected through documents. The clustering of the graph works in much the same way: more connected nodes are drawn together, but I would suggest that using heterogeneous elements in this way, the *formatting* of the documents is brought into relief: the clustering can now either be attributed to changes in the climate change discussions *or* changes in the way notes are taken. The later possibility is something Venturini draws the reader's attention to later in the paper in relation to a different visualisation.[87]

So Digital Methods approaches, I argue, can be leveraged to locate and study controversies on particular platforms *both* in the sense of studying the controversy and possible interventions of the platform. However, to distinguish interesting platform effects or interventions in the controversy from uninteresting ones, it helps to use relational and hybrid approaches to move past frequency to find patterns in more than just the popular or trending content. Yet this only works to the extent that one can qualitatively interrogate the content of, in this case, the tweets, but this practice is not specified or, more accurately, Digital Methods folds the sometimes implicit qualitative analysis into the final data visualisations (Kitchen 2014) which, just as with Mapping Controversies are an end goal in themselves.

Qualitative Analysis: Actor Worlds, and Rhetorical Strategies

Because, as I have argued, the role of participatory platforms in controversies warrants more detailed analysis 'in action' it seems important to define better the role of qualitative techniques in this work. Virtual Ethnography provides some inspiration in this regard but I think something additional is needed to grasp both how platforms represent a controversy and the arrangements and infrastructure through which they

[87] I also want to highlight this example because it shows that researchers associated with Mapping Controversies often turn their analyses back on the devices which enable them to map, just as Digital Methods researchers are both interested in mapping and studying platforms, though this tension is rarely made explicit.

are produced. It is important to remember that the classic study *Mapping the Dynamics* (Callon et al., 1986) outlined a full methodological approach which included not only co-word maps but also ethnographic studies and the literary analysis of texts. According to Law's chapter in the book, co-word is grounded in a form of qualitative analysis which proposes that words '…index forces which restructure the environment of science and technology in a variety of ways'(1986: 68). Scientific texts are seen as performative propositions for what they call 'actor-worlds': in the sense that they propose present or future configurations of heterogeneous (human and non-human) actors which may be realised in the real world to the extent that other actors and texts accept the 'translations' proposed. When these propositions take on a material instantiation, they become harder, more costly to unmake (Latour, 1991).

To give a related example of a similar approach to texts: in his study of Pasteur (1988) Latour suggests that an application of semiotic theory to historical, textual documents is one direction that STS can pursue given the place-bound limits of ethnographies in laboratories.[88] This 'socio-semiotic' analysis allows an ethnographic sensibility to be extended to larger socio-political controversies. Latour examined a science publication at the time of Pasteur and described the shifting configurations of actors (scientists, governments, farms, equipment, bacteria) these texts proposed. This would be a very ANT way of approaching online texts, in that it does not assume interests or claim to reveal latent meanings in text but merely detail the actors included (and progressively excluded) and their shifting relationships over time. This technique could be used as a more empiricist, less assumption-laden way to approach 'media framing' discussed in the last chapter without presuming the stability of particular articulations of a controversy.

This technique was *mainly*, but not exclusively, used to tell the story of Pasteur, that is the substantive content of the texts, but it could also be used to tell the story of diverging journal coverage of his accomplishments, that is take on a device-centred perspective. Partly this would be to link repeated configurations of actor-worlds (Kitzinger, 2000) to technologies, practices and 'backstage phenomena' such as the production of texts. But to dig deeper into how platforms as socio-technical devices, are

[88] A socio-semiotic approach combines the analysis of texts with an anthropological sensibility. Latour draws on Greimas's semiotics to analyse popular science articles of the time period. He treats the signs of the text as actors and charts their rise and fall to describe how Pasteur mobilised heterogeneous resources and crossed boundaries between science and politics to achieve his goals.

implicated in the production of texts, they must also be read in another way – not just for their content but in terms of the resources marshalled in support of that content.

Latour and Bastide (1986) outline one such way of reading of scientific texts, as if they were literature, by outlining various rhetorical strategies authors use to strengthen facts and channel the reader through the article so that their objections are anticipated and closed down. This technique, which is elaborated in (Latour, 1987) considers which resources and technologies scientists invoke in order to ensure that their facts 'win' over other competing facts, something which would also be applicable to news articles, blogs and Wikipedia pages, all of which have their own unique ways of legitimising propositions. Claims about a controversy may be strengthened with a particularly official tone or may be attributed to experts or statistics or, in particular, other online texts.

Latour especially talks about how scientific articles are reinforced through references to other articles and how these references are 'modalised' (qualified by positioning texts as supporting their cause, questioning them or even playing texts off each other). Bealeau (2005) relates Latour's study of modalities to the qualification of hyperlinks in virtual ethnography – studying the surrounding context (placement on the page) and the subtle ways in which the links are framed: positive negative ambivalent, etc.

But as discussed earlier, in contrast to science articles, many of the references in a news story, a blog or a tweet are *implicit*, not designated by a material hyperlink. The advantage of adding ethnographic, open-ended tracings is that they do not necessarily rely on hyperlinks. Virtual ethnographers can jump to websites which are implicitly referred to by websites or discussed by informants in interviews. These are alternative ways of jumping between research sites, rather than device-determined tracings. In his analysis of approaches to 'multi-sited ethnography', Marcus (1995) outlines several logics to linking previously disconnected research sites, including following actors, following objects and following symbols which all present possibilities for tracing. But one logic which may be particularly relevant to controversies is 'following the plot' – 'Reading for the plot and then testing this against the reality of ethnographic investigation that constructs its sites according to a compelling narrative…' (Marcus, 1995: 109). In other words tracing 'links' might mean checking up on the actor-words proposed by texts and determining if the actors accept the translations proposed or see if how propositions have been modified. If a Twitter user makes a simple proposition, it

can be traced back to a news source it links to and the news source can be traced back to its (implicit) or (explicit) sources and so on. This may even be traced back to invoked 'eyewitnesses' or offline sources. The purpose of this process would not be to locate a definitive truth of the controversy but rather to follow the chain of transformations as particular propositions or associations between actors are modified and translated between platforms.

04. DECOUPLING METHODS FROM PLATFORMS

Armed with Digital Methods tools and STS inspired forms of textual analysis should be enough to study the role of participatory platforms in controversies 'in action'. But in this final section I want to raise some further concerns about the possible tension between platform-specific digital methods and non-platform specific controversies. I am going to argue that while our methods, both digital tools and qualitative approaches are necessarily entangled with online platforms, following controversies might entail in subtle ways decoupling the two. While earlier I took as granted the seeming fit between methods or devices of the platform and our social science methods, here I want to explore the tension.

The problem I am pointing to is that Digital Methods and in fact most digital tools are by design linked to easily calculable and scrapable data (hyperlinks, hashtags, articles etc.) when the digital traces which may interest researchers and be relevant to the particular controversy (as opposed to the platform) may stubbornly resist this logic. Forces or negotiations may lie in backstage phenomena which less easily scrapable or only through other platforms or devices altogether.

In particular, it should be said that a lot of the easily usable data is expressed *quantitatively* on these platforms, through frequencies and rankings. This potentially makes for a somewhat problematic division of labour between 'quant' and 'qual' in which the large scale maps and the micro-sociological analysis have jurisdiction over different sorts of materials. The main problem however is that these jurisdictions are largely determined by *the platform* and the API, not the *controversy*. In this last section I will point to how we can (partially) resist this division of labour.

Redistributing 'quant' and 'qual'

Firstly it is important to recognize that historical debates about quantitative and qualitative methods, which I do not have space to rehearse, are generally about their *capacities*, their pros and cons. For example, quantitative methods give wider scope while qualitative techniques give nuance, depth and context to larger trends (Fielding & Fielding, 2008).[89] But if methods are entangled with objects and devices being studied, then the balance of quantitative and qualitative techniques should to some extent be empirically determined. Venturini (2015) offers a helpful distinction between 'digital traces' and 'digital data': digital traces are any information stored in bits on servers whereas digital data have in some sense been readied or formatted for analysis (not necessarily in the way desired by social scientists though).[90] Likes on Facebook are easily quantifiable whereas unformatted text strings, images, audio etc. take more labour and have more costs to being made analysable in large scale, automated ways.[91] There are also of course countless offline practices, traces on other platforms and also psychological and affective states, which may only be alluded to through existing traces, or not be traceable at all.

These data-structures encourage the selection of quantitative methods but also qualitative techniques. Facebook's networked system of profiles seems to suggest classic sociological techniques like social network analysis (Hogan, 2010) where as the short call and response conversations on twitter might call to mind Goffman-esque approaches to talk (Murthy, 2013) with the effect that quantitative and qualitative researchers perform different ontologies when it seems that the phenomena we are researching sit somewhere in between. The difficulty is that these silences and gaps in analysis may be *patterned* (Kitchen 2014) and mask phenomena crucial to a controversy. In past studies of participatory settings, it is the exclusions – which ways of speaking or which actors are excluded from the proceedings – which is as important as what is included. While researchers are to a large extent dependent on devices like

[89] Mixed methods have often been used to 'validate' each other, to converge on a reality independent of the methods employed but researchers who are skeptical of this claim still see 'quant' and 'qual' as contributing to a depth of understanding (Fielding & Fielding, 2008).
[90] As stated earlier, this is not to say that digital traces are 'raw', they are already formatted in the act of collection, but some data are more readily available for accepted modes of analysis.
[91] The danger if we only adopt the methods afforded by the platform, then the performative character becomes assumed: the platform might loom too large in the analysis and we might miss the moments of resistance and break down. This is not to be confused with a tension between the technical versus the social, scripts are both human and technical and resistance to them may be equally technologically equipped.

online platforms for access to the study, the idea of following controversies suggests that they might, temporarily at least, destabilize and unsettle the *modus operandi* or the frame of a device and this might mean tracing, in Callon's language, the overflows which overwhelm the frame.

For help with this problem we can, I suggest, draw inspiration the work of Gabrielle Tarde, as read through ANT. In the preceding discussion, Tarde's use of monads was helpful for thinking through a basic ontology of associations in the web, but perhaps in the case of online platforms other aspects of his work may be relevant.[92] In another attempt to revive Tarde and make him amenable to STS, Latour and Lépinay (2009) describe his unique version of economics. They describe how Tarde criticized the neoclassical economists not, as one might expect, for trying quantify the unquantifiable world of associations and interpersonal interactions, as economic anthropologists would have it, he argued that they were quantifying the wrong things.

Tarde proposed that there were other phenomena which pertained to the exchange of goods that were not captured by the sort of measures economists devised. These were the 'psychological' factors, by which he meant not inner mental states but moments of *intersubjectivity* when monads imitated or innovated off other behaviours and characteristics – such as when something was valued as true, as beautiful etc. Counter-intuitively, Tarde proposed that these can and should be *quantified*. Beliefs and desires in the sense of mental selections or valuations of objects are quantifiable (how many believe in something) but sensations, which are intermingled with beliefs and desires are qualitative and difficult to enumerate.

However this did not mean dismissing the economic measures. In much the same way that digital methods practitioners see the formatting work of social media metrics, he saw economic measures (which he called value metres) as making possible calculative situations. Tarde saw these devices, to place this into Callon's language, as performative – in the sense that they did not so much describe the economy as format it, and thus are a topic for analysis, not sitting outside of it. Thus Didier (2010) explains how he analysed crime statistics, while simultaneously reflecting on how the re-

[92] Here I say inspired by because there are many ways of understanding Tarde, Tarde the crowd theorist (Borch, 2012), Tarde the reactionary (Toscano, 2007), so I will confine this discussion to those aspects of Tarde which have been taken up by STS theorists rather than presuming to speak to what he actually meant. My interpretation is also limited, at this time, to texts which have been translated into English.

categorisation of criminality in his data set influenced the likelihood of re-offence – he was analysing the object (crime) and the formatting of the object through methods (crime statistics) which is the sort of two pronged analysis I have put forward in this thesis: to study both particular objects and the devices that mediate them.

Although I will not attempt anything as ambitious as calculating belief and desire, I will take this as a call to quantify those traces which are less obviously calculable while seeing the easily calculable ones in relation to the 'value meters' of participatory media (likes, rankings , shares etc).[93] I will propose that the capacities of 'quant' and 'qual' need to be redistributed: we need to *qualify* these quantitative traces (study *how* likes and hashtags are shared, not just *how many*) and make the incalculable more *calculable* – find ways of visualizing text and other unstructured data, when the controversy demands it. Co-word represents one way of moving past quantitative, frequency based traces, but I will propose others. This still however means working with the limitations and affordances of digital traces. Since particular devices mediate controversies and issues, we cannot escape them but following controversies might suggest *different devices* which are less easily analyzable from the perspective of quantitative methods or platform-specific approaches.

This also might present an opportunity, if we are less beholden to dominant data structures, to analyse these devices in different ways. One of the capacities of digital devices, which is not always capitalised on in sociological analysis, according to Law and Ruppert (2013) is continuous time. So rather than dealing with fixed time slices, there is the potential for more granular analyses which reveal the rhythms and intensities of phenomena – something which Tarde also valued in his version of statistics and something which I argued earlier is essential to studying controversies 'in action'. Law and Ruppert also point out that digital data is conducive to *visual*

[93] But what would this look like? Didier, in describing Tarde's view of statistics, points out that Tarde was not interested in enumerating categories and classes but looking at changes over time. He saw society as a kind of contagion in which ideas passed from monad to monad but were changed through acts of imitation and innovation. The craft of statistics was to find something which could be held constant which would then allow the researcher to identify variations. Desires and beliefs, were seen as perfectly similar but they were coloured by sensations which were not. Things like voting patterns and the stock exchange and marriage rates could be seen as rays or beams of imitation. They are the same thing but perceived differently through individual sensations. In the case of participatory media, one could take 'likes' or posts or shares, the calculative traces of participation as the visible results of the psychological act of choosing to get involved or resist or innovate in relation to others.

analysis rather than the end result being numbers.[94] There is currently a movement in visual sociology, which classically deals with photography and film, which argues to employ more experimental modes of visual analysis (Grady 2007). Michael Guggenheim (2015 forthcoming) describes how the various media through which sociology represents its objects have classically been statistics and fieldnotes and transcripts which are highly abstracted from their source. This is the case, he argues, because of a misguided attempt to imitate science – they equate objectivity with being free from manipulations. But using an STS view of scientific representations, Guggenheim shows that the visual materials of natural scientists is, in practice, characterised by lots of manipulations. But rather than massive leaps, for example, from the text of survey questions to scatterplots of responses, they are chains of what he calls 'tight' translations (in which every step in the sequence can be accounted for).

This work, which draws on the sociology of scientific representations (Latour, 1999; Lynch, 1988), draws attention to the constructed, or rather the performed, nature of visibility. Michael Lynch in his contribution to the volume *Representation in Scientific Practice* (1991) describes the twin processes of selection and mathematisation which accompany the production of visual materials in science. Lynch emphasises that selection is not best conceptualised as a passive process of filtering but as an active practice of drawing out certain characteristics. Mathematicization is the process of then imposing geometry and edges onto visual materials such as photographs and imaging outputs with the aim of making them amenable to measuring and mathematical operations. While natural scientists often capitalise on the perceived geometry of different specimens, i.e. non-human nature participates in the 'discovery process' by offering certain contours and edges to build on, scientists may also actively produce environments and settings to turn nature into inscriptions. Lynch gives the example of placing lizards in an enclosure actually containing a grid in which their winding movements can be captured as the crossing of thresholds. So there is a tension within scientific work between using what is already there as anchor and imposing particular frameworks on the proceedings.

Although many of the new platforms which feature in this thesis are associated with a regime of 'openness', where all data is made public mentioned earlier, Tkacz reminds

[94] Andrew Barry describes possible ways that Tarde's program might have been realised and specifically gives the example of French linguist Rousselot who used creative methods like a facial armature to measure variations in pronunciation (Barry, 2010).

us that (2014) openness actually conceals as much as it reveals. Or as I will discuss in the next chapter making something visible means actively using equipment or processes or manipulations, *both* in terms of those methods and metrics built into the platforms but also in terms of our own work on top of this. I will refer to visibility as this process of selectively highlighting phenomena whether in the case of methods in platforms or our own social science methods. When researchers align their methods with the methods, value-meters and metrics of the platforms they are following the contours of the device, but if the device alone is not the object this might require more, not less manipulation.

What I have just proposed is that one way of escaping the pull of platforms is to rethink the division between quantitative and qualitative techniques: *qualify the quantitative and make the incalculable more calculable*. I also suggested that modes of visual analysis which focused on time were particularly appropriate for transactional, digital data and could offer a more exploratory, interpretive mode of analysis. This may sound abstract at this point, but the idea will become more clear over the course of the empirical chapters which will explore different ways in which platforms may restrict the analysis of controversies: through boundaries and silences, the separation of number and text and content and material infrastructure.

In each chapter, I will study a particular controversy or series of controversies using digital methods techniques to identify potential starting points on different platforms. Then I will focus on how the controversy is read through a particular platform, by viewing the platform as a series of device in a shifting ensemble of devices. Then, through an iterative qualitative and quantitative process I will develop new types of data visualisations which both visualise the controversy and can be reflexively analysed to apprehend some of the politics of visibility.

05. CONCLUSION

In the last chapter I argued for a decentring of the object of study from platforms using controversies. In this chapter I considered how methods are entangled with these platforms and how this potentially enables exciting new forms of research but also how these platforms might constrict or direct the study. I first argued that so called big data presents an opportunity to rethink the relationship between quantitative and qualitative methods and looked back at successful configurations of methods focused

on hyperlinks: virtual ethnography and hyperlink analysis. However these techniques, while still relevant to large sections of the web, are less effective for online platforms. I also described how Digital Methods, have successfully repurposed platform metrics in a similar way but there is yet no parallel qualitative approach. I offered an STS informed reading of texts: not only looking at rhetorical strategies – how claims are advanced and strengthened – but also following texts outwards to networks, actors and other texts beyond the particular article. This could even lead to particular informants or offline field sites. However, while the methodology proposed above and the visualisations I produce in the following chapters represent a way of studying controversies which could lead to offline encounters and I have in fact conducted interviews and produced fieldnotes in ethnographic locations, there is not space in this study to go into detail about these. For the sake of clarity, I will confine my methodological discussions to the online component of my analysis though I will return to the offline in the conclusion.

In the case of these STS informed methods, both qualitative and quantitative techniques can both be used to locate events or controversies but also be reflexively turned back on the platforms themselves to reveal how they intervene in and format them. But I cautioned that in order to study controversies through platforms, not only platforms themselves we may need to move beyond the low hanging fruit of digital data and seek to map less easily calculable *digital traces* as well. This is important because traceability and quantification are not just affordances of platforms which make automated forms of research easier – they are integral components of these platforms which may potentially shape controversies. As will become clear in the subsequent chapters the variable visibility of certain digital traces is actually quite important to how controversies play out.

It should be said the empirical studies in this thesis are pilot studies and they represent a progression towards what I have detailed in this chapter, which is dependent on both platforms (the traceability of phenomena) and resources (the help of programmers). In each subsequent chapter, the quality and accessibility of the data increases as the platforms become more calculative and this allows for more granular analyses and closer combinations of 'quant' and 'qual', culminating in Chapter VI. The chapters that follow will go into much more empirical detail than is perhaps normal for a methodological argument, but my gambit is that methods, even semi-automated

mapping techniques need to emerge in dialogue with specific empirical objects and problems rather than be imposed on them by the affordances of media devices.

IV. BOUNDARY WORK: WIKIPEDIA AND INDETERMINATE SETTINGS

It has long been accepted within STS that domains of social life have indeterminate boundaries. Taken for granted divisions, such as those between science and media or experts and laypeople are not given but an accomplishment of social and technical arrangements. Methodologically this means 'following the actors', or 'tracing the overspills', which constantly criss-cross these institutional and conceptual domains, while analysing the on-going maintenance of boundaries as a topic.

This first empirical chapter starts from the observation that online platforms, when understood as socio-technical devices, are no different: information, actors and infrastructure all pass between these websites. However platforms *also* feature boundaries of a more *practical* sort, enforced through the structuring and availability of various data, which restrict the mobility of either qualitative or quantitative analysis. Yet these more technical boundaries such as the walls separating online platforms or standing between front and backstage areas and different data formats, are just as permeable and socio-technically constituted – research subjects and devices cross them in practice all the time. The argument of this chapter is that sometimes these practical / technical boundaries must be questioned in order to question the conceptual ones.

One supposed promise of new online platforms is that they can upset the normal hierarchy of information flows between science, media and audiences. This is supposedly due to the value of 'openness' – the idea that the transparency of information will allow for a marketplace of ideas – and the best ideas will rise to the top (Tkacz, 2014). Knowledge can be organised (seemingly) without recourse to credentials or hierarchy. This is at least one of the premises of one of the quintessential Web 2.0 platforms, Wikipedia, a collaboratively written encyclopedia which can be edited by anyone. Wikipedia's 'Wiki' software makes available every version of every article along with comments and extensive forum-style discussions.

Open platforms present researchers with a methodological problem: the amount of data produced is staggering, too much to process by normal methods and both researchers and platform participants must be technologically equipped to make sense of it. Yet, remembering Venturini's (Venturini et al., 2014) distinction between digital traces (anything stored digitally) and digital data (pre-formatted for analysis) – this means that certain phenomena will be harder to access and analyze than others. However, following controversies, and not just platforms, demands that one must analyze what is most consequential for the controversy, not just what is easiest to capture.

This chapter analyses coverage of the Fukushima disaster through the English-language version of Wikipedia. From the first suggestion of a nuclear incident, a largely anonymous collection of editors wrote an extensive article as the situation unfolded with up-to-the-minute information on this far-flung event culled from media reports. What is interesting about this incident is that it undermines the presumed flow in disasters from science (or experts) to journalists to audiences. Anonymous editors on Wikipedia, who may themselves be journalists, affected citizens or nuclear scientists, bypassed the normal channels to obtain primary information and question mainstream media accounts. Yet in order to describe this crossing of domains requires cutting across different types of data structures with different analytic capacities.

Before confronting these methodological challenges, I will start by discussing literature on what has been called 'boundary work' (Gieryn, 1983, 1999) between science and media and why online platforms may redistribute these relationships. I will then analyse the Wikipedia article and talk about Wikipedia as a socio-technical device before diving into a more detailed analyses of this particular controversy on Wikipedia,

01. SCIENCE AND THE MEDIA

As discussed in Chapter II, researchers studying the intersection of science and media often rely on taken-for-granted models of how knowledge trickles down through society. There is, for example, the so-called dominant view in science communications, that scientists produce authentic knowledge, which must be simplified (or possibly distorted) and communicated to the public 'downstream' through the media (Hilgartner, 1990). Studies of mass media have also been premised on a 'gatekeeper'

(White, 1950) model in which journalists select what is disseminated to (passive) audiences as news. Both of these models, which purport to describe different ends of the chain from science to media to the audience, have been challenged theoretically and subjected to numerous empirical counter examples; yet they have rarely been interrogated *together* as part of the same process.

In a well known discussion paper, Stephen Hilgartner (1990) demonstrates, using the case of a paper on cancer prevention, already geared toward an 'interested lay audience', that any way of establishing a clear boundary between pure scientific knowledge and simplified public-friendly versions is untenable.

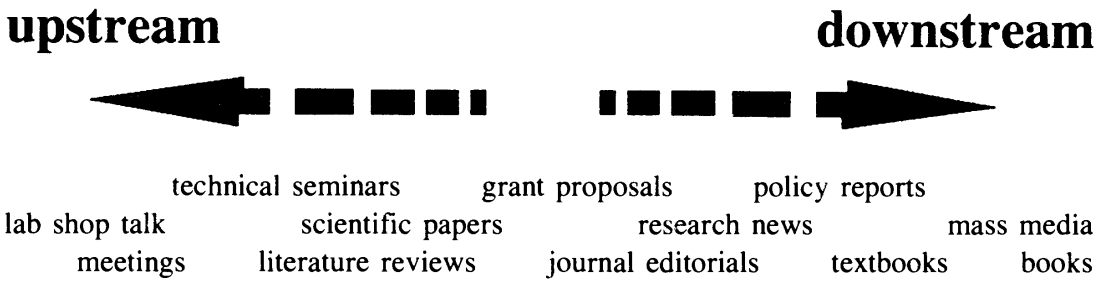


Image 1. From Hilgartner 1990

This is not to say that knowledge is not geared towards more specialized or more generalized contexts, but it is a spectrum full of overlapping and permeable categories not a binary opposition. (see Image 1). However, only scientists are allowed to use this model to distinguish 'appropriate simplification' from damaging 'distortions' in knowledge, when it suits them.

A classic example of this was the cold fusion scandal, alluded to briefly in Chapter II. There are numerous detailed accounts, so I will only remind the reader of the basics. Pons and Flieschmann, two electro-chemists, held a press conference to announce that they had achieved the holy grail of nuclear fusion at room temperature using a relatively simple laboratory set up. Pivotally, they announced this before they were vetted by the peer review process in a respected journal, which allowed them to bypass a priority dispute (Merton, 1957) with a neighbouring university and directly enrol government actors and funding bodies. After the initial furore this caused, experimental refutations and counterclaims began to pour in, culminating in another

press conference in which the claims were resoundingly denounced by the scientific community.

While some scholars have focused on the discursive strategies of the debate *within* science (Collins and Pinch, 1998; Simon, 2001), several others (Lewenstein, 1995a; Lievrouw, 1990) have examined the abnormal presence of the media in the event and the (temporary) reversal of the usual order of knowledge dissemination. Gieryn (1999) describes the event in terms of 'boundary work', that is battles over what counts as proper science. In the first press conference, the scientists expand the possible contexts in which scientific discovery can happen by inviting the media to 'participate' in the discovery. But the second press conference was used by other scientists to expel the press (and Pons and Flieschmans's claims) outside of the bounds of true science.

The many commentaries from STS scholars, naturally do a thorough job of opening up the normal processes of science to scrutiny, but I would argue that many, such as Gieryn's leave *the media* and their processes largely black boxed. They become passive transmitters of the scientists' messages.[95] Trevor Pinch even goes so far as to argue that while the presence of the media was abnormal, the controversy played out in a way sociology of science was equipped to handle, (Pinch, 1994) but Lewenstein (1995), who I discussed earlier, argued that the media at least affected the pace and dynamics of the controversy if not the economic and social framing of it as well.

But to properly accommodate the media into the controversy, one has to confront another set of tired assumptions about the proper dissemination of knowledge. Sender-receiver models of media have been critiqued for much longer in media literatures, firstly as a 'two-step' model (Katz and Lazarsfeld, 1955), then the process of both knowledge production and reception was opened up as 'encoding / decoding' (Hall et al., 1980) which was then elaborated as more of an iterative circuit in which audiences in various ways feed back to or are anticipated by media (Miller et al., 1998). But in most formulations, professional journalists and editors are in some respect 'gatekeepers' (White, 1950) who determine which information is allowed to reach the public or audience in a similar way as scientists (supposedly) manage the flow of information to the media.

[95] It might also be argued that Gieryn's understanding of boundary work is more cultural and grounded in discursive exchanges, underestimating the role of objects and technologies in enforcing boundaries.

Even though sender-receiver understandings of media has long been questioned, two types of empirical phenomena have further complicated these views. Firstly literature about the advent of new media and later social media has focused on the potential of platforms like blogs and p2p websites to allow everyday people to contest mainstream media messages. The boundaries between the (passive or not) audience and active content contributors have become hopelessly blurred. Wikipedia, the focus of this chapter, in particular has been crucial in the conceptualization of 'gatewatching' (Bruns, 2005), a concept which describes how non-professional media actors are now able to 'curate' or select alternative conceptions of what is news out of a wider array of information sources. Yet even gatewatching suggests that social media actors are *secondary* to mainstream media, re-framing and rearranging content, but only after the mainstream media has made their selections. As I hope to make clear, these relationships between mainstream and 'alternative' (Lievrouw, 2011) media platforms are far more complex than these models allow for.

Disasters

There are also certain kinds of events, such as unforeseen disasters, which further allow for the possibility of participatory media actors crossing the media-audience boundary. It has always been the case that disasters de-stabilise normal journalistic working practices (Molotch and Lester, 1974) and may require them to engage with citizens on the ground (Sood et al., 1987) but social media, it is claimed, gives new powers to the non-journalist. Murthy and Longwell (2013) give several examples of disasters and accidents in which Twitter users were able to report from the ground before the mainstream media arrived on the scene. In the case of the Hudson River plane crash, a lone user with a phone was able to report the incident first. The media in these situations are merely playing catch up, attempting to vet these social media reports.

However disasters do not always have the same effect on jurisdictions between science and the media. It is hard to compare disasters, which are almost by definition unique and unpredictable, but one could argue that in socio-technical disasters (there are no purely 'natural' ones Hilgartner, 2007), in which spokespeople like scientists, experts, bureaucrats and engineers must speak on behalf of nature, techno-science and society,

that the boundary between science and media, far from being undermined becomes stronger than ever.

For example, Farías (2014) gives an account of a disaster in Chile in which an early warning system failed to predict a tsunami and make the call for an evacuation, resulting in the deaths of 1000s. Despite massive uncertainty and conflicting claims, bureaucratic procedures ensured that the experts in charge of detecting tsunamis would decide if a tsunami was probable before another set of experts would alert the populace and the media. Although this was a communication failure between domains of science (which appropriately enough fell along nature-society boundaries), the priority of scientific claims before mainstream media reports was forcefully maintained.[96]

Another disaster in which the science-media boundary was (largely) maintained is Chernobyl. According to (Luke 1987), there was a complete media-blackout after the reactor exploded and the city of Pripyat was quickly evacuated, but the Soviet government later changed tack as part of the 'glasnost' policy of openness. Once the information began to trickle out, however, the mainstream media had difficulty in reporting the highly technical radiation data – often confusing micro- for mili- Sieverts. The press seemed sceptical of the nuclear scientist's efforts to put the numbers into perspective, such as comparisons with background radiation, but did not have the expertise to contest them (Otway et al 1987). In this case the downstream model, in which official (expert) reports precede media interpretations (framed as distortions) before arriving at voiceless publics and lay audiences, was seemingly maintained, though it is no doubt more messy in practice.[97]

In contrast, however, the unfolding of the Fukushima nuclear accident, which bears obvious similarities, differed for two reasons. Firstly, partly due to the tarnished reputation of the nuclear industry, plant operator TEPCO was obliged by industry regulations to make gestures to transparency which occurred through a series of up to

[96] Now, to clarify, not all disasters have a high science content and rarely involve experimental scientists in laboratories as the inevitable fount of facts; it is rather engineers and technicians who are the proprietors of knowledge, nonetheless presented in scientific ways. Some experts arrive only after the fact in the form of audits and investigations (Hilgartner, 2007). Also, these disasters draw in other actors besides science and media, including governments and the private sector who may strategically enforce a controlled dissemination of knowledge.
[97] The above accounts, which are based on a review of mainstream stories may presume, rather than demonstrate, the efficacy of the upstream-downstream model.

the minute press releases and a webcam of the site. Secondly, Japanese citizens had the technology to pool on the ground information such as Geiger counter readings into alternative maps (Plantin, 2011) which could contest official figures through blogs and social media. Rather than waiting for carefully processed official reports or intrepid journalists – this partial and competing information could be digested, vetted or discussed in real time, as the controversy happened through various online platforms and news sites.

But of course all this circulating information both official and unofficial did not lay the groundwork for a kind of consensus about the event, if anything it made things more controversial. Despite TEPCO's supposed transparency, there was much outrage from Japanese citizens about the controlled dissemination of information including the (accidental or not) loss of temperature data at a key moment in the accident (Slater et al., 2012).[98] As Andrew Barry notes, making more information available in the name of transparency can actually create more controversies, because there will always be more that could be revealed (Barry, 2013). So in a situation of radical uncertainty, how is reliable information identified when the usual institutions (news and science) are in a reactive mode themselves? The Fukushima disaster and controversies over its coverage, seem like an ideal object with which to revisit debates about the relationship between science, media and audiences because it appears to destabilize both together.

02. CONTROVERSY AND DEVICES

One way of unthinking the dominant model of science-media-audience relations is to approach these types of cases through a particular controversy (see for example Lewenstein, 1995a) which means maintaining uncertainty with respect to the object of study, not deciding in advance what the controversy is, but instead following various actors attempts to settle it. This also means not deciding in advance if experts, media actors, platform technologies or lay-publics are most consequential for defining the outcome: these identities, who counts as an expert are also at stake in the controversy. Since this is the first empirical chapter I will take the reader successively through the

[98] According to Asahi.com http://ajw.asahi.com/article/0311disaster/fukushima/AJ201211200029 (Accessed 12 August 2015).

process of identifying the controversy through online mediations of it and then studying the production of these particular articulations through socio-technical devices.

Articulating Fukushima

First, how is Fukushima, an object of controversy, mediated by various online texts, platforms and technologies? If one knew nothing about the event in question and were to type 'Fukushima Disaster' into Google, the first two results, depending of course on your past searches and what the latest news is (all of which factor into how Google personalises your search) would likely be an article on the site WorldNuclearNews.org, published by the World Nuclear Organization, and an article on Wikipedia. Contrasting these two texts can help bring into relief what is at stake in such articulations of the controversy. This involves reading them both for their very partial presentations of 'actor-worlds' and the rhetorical and literary techniques being deployed, as discussed in the previous chapter.[99]

The Wikipedia page[100] 'Fukushima Daiichi-Nuclear Disaster' is a 17,000-word recap of the disaster. The reader first learns the basics: that the disaster was caused by an earthquake and Tsunami and that it registered level 7 on the INES nuclear incident scale which places it on the same level as Chernobyl in terms of severity. The introduction describes the multiple core meltdowns and hydrogen explosions but then moves into various speculations about health effects, in particular thyroid cancer and the continuing radiation leaks. These possible effects are currently given weight by their placement at the front of the article.

The body of the article starts with safety concerns about the plant, highlighting energy company and plant manager TEPCO falsifying safety records and ignoring tsunami warnings. This sets up TEPCO as the villain of the story. What follows is a forensic detailing of the events in each of the three affected reactors, which also get their own Wikipedia articles. This detail and references to reliable sources seem to accumulate around sites of controversy like scabs protecting a wound. For example the

[99] Of course what I am sidestepping is a discussion of how Google frames the debate through these page rankings, and particularly its symbiotic relationship with Wikipedia. This has already been dealt with extensively by Rogers (Rogers, 2004).

[100] The particular version I analysed was live on 16 October 2013, available here: http://en.wikipedia.org/w/index.php?title=Fukushima_Daiichi_nuclear_disaster&oldid=577437991 (Accessed 12 August 2015)

controversial assertion that plant manufacturer General Electric (GE) was warned about possible design flaws, requires not one but three references.

The style is an odd mix between the detached, dispassionate tone of an encyclopaedia article and a more journalistic urgency with the occasional editorializing line:

> Government agencies and TEPCO were thoroughly unprepared on almost every level for the 'cascading nuclear disaster' which was caused, in part, by a public myth of 'absolute safety' that nuclear power proponents had nurtured over decades.

But while the article is seemingly 'just stating the facts', this particular configuration of actors, institutions, fuel rods, valves and counters and regulatory documents presents the disaster as a result of incompetence by TEPCO in the design, sighting and later, disaster communications. Though this is the collective editorial line of the article, this conclusion does not appear to be the product of a clearly pro or anti nuclear agenda because it considers material from both sides (the article is damning on the safety of the plant but cautious about the attribution of radiation dangers). So while it has a clear argument it is self-consciously presented as 'neutral' at least in the journalistic sense of showing both sides of a debate.

The WNN article[101] (13,000 words) describes, in the style of an official report (complete with executive summary) the extreme events which precipitated the nuclear disaster. The article clearly spells out the amount of radiation released and the meltdowns but in contrast they make clear that:

> 'There have been no deaths or cases of radiation sickness from the nuclear accident, but over 100,000 people had to be evacuated from their homes to ensure this. Government nervousness delays their return.

This temporal framing of the event to include the past and present, but *not* the future, precludes the possibility of radiation deaths or cancer being attributed to the disaster many years from now, which is of course the primary concern for anti-nuclear activists. The statement is also interesting because the Japanese government emerges as a

[101] Available at: https://web.archive.org/web/20131013065805/http://www.world-nuclear.org/info/Safety-and-Security/Safety-of-Plants/Fukushima-Accident/#.U7_DDY1dVz0 (Accessed 12 August 2015)

separate, potentially culpable player, where in the Wikipedia article, TEPCO and the government are almost conjoined entities. The text details many of the same objects as the Wikipedia article but noticeably does not mention TEPCO's falsification of information. It offers that TEPCO designed and sited the plant based on 'the best science at the time', anticipating earthquakes of a reasonable magnitude. In this articulation, the disaster was not a failure of judgment and design, but sufficient and reasonable assumptions being overwhelmed by an unruly 'nature'. By this logic, if risks are calculated correctly, then accidents beyond the remit of reasonable calculability are anomalies – acceptable collateral damage of a desirable energy source.[102] This is also a very narrow framing of the event without a sense of wider societal context or even debates about nuclear – which are placed out of the frame.

This brief analysis of what I described as the actor-world of the texts (the entities present and their relationships) and some rhetorical techniques, shows how important the subtle differences of selection and emphasis are for advancing alternate realities of the event. One article places the balance of explanation on the side of society (human error) while the other attributes the event mostly to nature (the tsunami). In some ways the coverage of Fukushima conforms to a classic understanding of knowledge controversies in which the battle lines are drawn around various facts and articulations of the controversy, but as we proceed the nature of these disputes becomes less about facts and more about 'sources'.

To stop at this point with the analysis of texts would be to rely on a, more or less, social constructivist account, in which the differences between the texts might be attributed to the 'interests' of the nuclear industry or the 'culture' of Wikipedia.[103] In order to go deeper it is important to ask questions about how these texts were produced out of particular socio-technical arrangements 'in action' through the unfolding of specific controversies. These may include some of the substantive controversies already mentioned in the literature over missing radiation data or the culpability of TEPCO or more media specific disputes over the writing of the article – in fact it is the interface between the two which interests me. The production of the WNN article is of course

[102] As Wynne notes, often risk is framed in terms of knowable or calculable risk as opposed to incalculable or inconceivable risk – unknown unknowns (Wynne, 2011)

[103] There is already an account of Fukushima on Wikipedia entitled 'Social Construction of Knowledge on Wikipedia' (Hara and Doney, 2015) which compares the Japanese and English versions of the Fukushima articles, attributing differences in content and editing behaviour to cultural differences.

mostly black boxed, with their editorial process behind closed doors – although one can get some insight by using the internet archive (archive.org) to see past versions of the article.

The process behind the Wikipedia article on the other hand is extensively documented. One has only to click the 'History' tab to see a list of thousands of versions of the article, including commentary by the editors, or click the 'Talk' tab to see an archived discussion forum used by the editors to settle disputes and ask questions. Wikipedia is also interesting for this study because of it's supposed openness and relation to expertise, as I will describe in the next section. I will first talk about Wikipedia as a socio-technical system but later will focus on specific technologies and devices within it.

Wikipedia's Device Frame: Managing Controversy

To understand Wikipedia as a device which frames and formats these media texts, it is helpful to start with some of the literature to get a sense of what one might call the *modus operandi* of the device. According to the literature, it might be said that Wikipedia as a device is intended to host and manage controversies, but such a device 'frame' is both accomplished socially though policy and routines and technically through bots, scripts and architectures and such attempts to contain or supress controversies will rarely hold.

Wikipedia was founded in 2001 after the failure of Nupedia, a more conventional online encyclopaedia written collaboratively by experts, which proceeded so agonisingly slow, it was eventually abandoned. Wikipedia in contrast, employed collaborative Wiki software and was, crucially, editable by anyone – in fact, all users are referred to as 'editors' – with no priority given to expertise or qualifications (Sanger, 2005). Wikipedia editors have produced nearly 5,000,000 articles and it is one of the top 5 most visited websites.[104]

Ethnographer Joseph Reagle (Reagle Jr, 2010) describes Wikipedia's key innovation as the doctrine of 'assume good faith' in conflicts. He gives the example of editors patiently listening to and dealing with a white supremacist editor until the dispute is resolved, despite major differences of opinion between the participants. Ethnographic

[104] Available from http://en.wikipedia.org/wiki/Wikipedia:About (Accessed 12 August 2015)

and historical approaches to Wikipedia posit a *cultural* explanation of Wikipedia as a particular type of community, arising out of free, libre and open source software (FLOSS) culture, stabilized by norms and shared goals (but certainly not without its internal tensions).[105] Without discounting these cultural explanations it is more in keeping with the ANT inspired approach of this thesis to ask how these cultures are manifested and maintained in material policy documents which are constantly invoked and discussed.[106]

In the page describing the central policy of 'Neutral Point of View' (NPOV) (which is also written collaboratively by the community), NPOV is defined as:

> …representing fairly, proportionately, and as far as possible without bias, all significant views that have been published by reliable sources. (Wikipedia Editors, 2013a) *Emphasis original.*

This means that Wikipedia will only represent what 'reliable' sources have claimed. If reliable sources disagree, then both positions must be presented as an open controversy. In an analysis of Wikipedia policy, Tkacz (Tkacz, 2012) explains that the central tenant of 'Neutral Point of View' (NPOV) was intended to allow editors with very different perspectives to collaborate without agreeing on what is true.

For Tkacz, NPOV sets up a two-fold relation to truth 1) a distancing of Wikipedia from truth battles in the world by focusing on what reliable sources say is the truth and 2) a set of criteria for selecting these sources in a 'neutral' way. So sources are crucial to the way Wikipedia functions. In fact the existence of at least two reliable sources on a topic is a prerequisite for any article being 'notable' enough for inclusion in the encyclopaedia. This already sets up a socio-technical boundary between reliable sources who make claims and Wikipedians who report them.

But there are further boundaries of reliability: Wikipedia's 'Verifiability' policy states that Wikipedia prefers third party (independent of the topic being covered) sources, published by institutions which, '…have a professional structure in place for checking

[105] This is not unlike accounting for the success of science in terms of a shared culture, norms and rewards (Merton, 1973) which to some extent may buy into scientists, own version of themselves.

[106] Following Entwistle and Slater (Entwistle and Slater, 2014) although ANT does not deal with culture as an explanatory tool, often research subjects will invoke culture in various ways as an object which they themselves materialize through various practices.

or analysing facts, legal issues, evidence and arguments' (Wikipedia Editors, 2013b). Scientific or peer reviewed publications are privileged but non-academic books by respected publishing houses and mainstream newspapers are also considered reliable.[107] So as a device, Wikipedia makes possible the participation of (potentially) anyone, regardless of credentials, by creating strict rules about the provenance of knowledge claims.

It will not have escaped the readers attention that this understanding of neutrality, that is the systematic balancing of differing views, might owe more to the journalistic usage of the term than something having to do with an encyclopedia. Indeed, many authors see Wikipedia as a potential model for how alternative, user-generated news sites can remain open to everyone and yet still provide (reasonably) accurate information (Bruns, 2005; Lievrouw, 2011; Thorsen, 2008).[108] It would be easy, coming from an STS perspective to think of Wikipedia in terms of knowledge production, like a laboratory or even in terms of open-source software development (Kelty, 2005) but as I argued in Chapter II, one of the ways of understanding the contribution of online platforms may be in relation to traditional media. As I will show, the Fukushima disaster brings out these tensions between encyclopaedic and media forms of knowledge as an empirical topic for the participants.

All of these combined policies are meant to *diffuse* controversies which may result from collaborations between editors with drastically different world views. Yet, these policies mean very little unless they are materially enforced. Firstly, when the neutrality or reliability of sources is questioned, users often invoke policy by placing a hyperlink to the policy page in their comment (e.g – NPOV). If the users disagree about the letter or even the spirit of the policy they can discuss it in the Talk page or take it to an administrator to settle (e.g Administrators Noticeboard) or if there is indeed an

[107] There are also numerous exceptions: blogs or self-published sources are frowned upon, except as a primary source when the article is about them (and qualified as such e.g. - 'Blog A claims that ___') or when blogs are affiliated with already vetted news organisations. There are finally also supporting policies like No Original Research and Synthesis which state that facts from one source cannot be combined with facts from another source to create new facts which do not exist in either. This presents a hierarchy of interpretation: scientist, or when none are available, journalists, can interpret facts but rarely bloggers and never Wikipedians.

[108] There has even been an explicit attempt, Wikinews, a Wikipedia side project, (see Allan, 2006; Bruns, 2006) to apply Wikipedia's editing process to breaking news, incorporating first hand accounts. Wikinews is still an active project but has failed to catch on in the way Wikipedia has, partly because editors interested in news will be more likely to invest time in the corresponding article on the larger and more prestigious Wikipedia project.

exception they can re-write the policy itself. Users who repeatedly ignore policies can be banned or suspended through a system of judicial hearings.

Technicity and Visibility

But how are these transgression and Wikipedia-specific controversies identified in the first place? Wikipedia's own narrative would try to position this accomplishment as a matter of 'openness', errors are corrected and bias is rooted out because everything is visible and documented. As Tkacz explains, this has its roots in Open Source FLOSS culture but also the political theory of Hayak and Popper (2014).

But I want to stress that this is also to an important extent a *technical* achievement (Niederer and van Dijck, 2010). Openness is not given: with four million articles there must be a technical infrastructure which allows editors to monitor certain activities, to direct their gaze. Openness and transparency must be actively produced; which is why I prefer to speak instead of *visibility*. Visibility, as I proposed in the last chapter is not open or closed but represents a spectrum; visibility is not constituted through an all-seeing panopticon only crude 'oligopticons' (Latour, 2005) which give partial, directed views.

For example, policy is policed through automated bots which scan all of Wikipedia for tell tale signs of joke edits and vandalism, though they frequently make mistakes. On the Fukushima Daiichi Nuclear Plant page, just as the news was reporting possible leaks of radioactive material, an anonymous edit appeared which wryly added 'Fresh Sushi has been made readily available for all staff involved.' This was removed seven minutes later by ClueBot NG

> ClueBot NG **Revision as of 11:46, 11 March 2011**
> *(Reverting possible vandalism by 99.247.26.252 to version by Shiftchange. False positive? Report it. Thanks, ClueBot NG. (344717) (Bot))*

ClueBot uses a database of classified edits to identify vandalism with a Baysean algorithm – not blacklisted words or heuristics. But ClueBot also ignores edits by users with a certain number of edits. This is a technical materialisation of a largely unspoken convention that experience matters on Wikipedia, even if expertise does not. So even if boundaries based on expertise are removed, more subtle ones may be created in their

place. A much simpler technology, but just as important, is 'watch lists' which allow users to monitor particular articles for changes made, like an RSS feed.[109] These allow editors to police novice users who may unwittingly change a page when it has been previously 'settled' or has achieved 'consensus' in Wikipedia terms.

Edit counts are frequently used by users to profile each other (are they inexperienced?) and tools like Wikichecker are used to monitor the editing history of users for the detection of bias. But similar metrics are often used to monitor the articles themselves. Frequently, Wikipedia editors discuss edit counts as a measure of their controversiality: that is articles which receive a lot of edits in a short period of time are seen as 'hot'.[110] The size of articles in bytes is also an important metric because Wikipedia has policies about the optimum size an article can be before it should be split into smaller articles. When articles expand out of control, this is often due to their controversial nature. Both of these metrics are often used in user created tools, some of which have been incorporated into the normal Wikipedia interface.[111]

These technologies, the logging of edits and the archiving of discussions produce a specific form of sociality in which users are aware they are being watched and that their record of edits may be used to judge their past actions. Users are also dissuaded from careless or improper edits because they can be easily found through bots and various tools. Wikipedia as a controversy management system is thus a joint achievement of human and non-human technologies and systems which often represent material stabilisations of norms and processes. These socio-technical arrangements create a sliding scale of visibility in which certain actions and areas of Wikipedia become more traceable and analysable than others – which has implications for participants as well as researchers as I will explain later.

The above discussion gives some sense of the complex interaction of policies, work cultures and technologies which form Wikipedia's *modus operandi*. In terms of devices, Wikipedia as a technologically-equipped, sprawling bureaucracy is more reminiscent

[109] RSS stands for Really Simple Syndication, a technology which allows subscribers to receive notifications when pages are updated or created.
[110] See for example: http://en.wikipedia.org/wiki/Wikipedia:Database_reports/Pages_with_the_most_revisions (Accessed 5 July 2015)
[111] Another set of metrics for monitoring Wikipedia pages: http://tools.wmflabs.org/xtools/articleinfo/index.php?article=Fukushima_Daiichi_nuclear_disaster&lang=en&wiki=wikipedia (Accessed 5 July 2015)

of the *dispositif*'s described by Michel Foucault, than the devices of Callon, and yet, examined through a contingent controversy, cracks start to appear in these imposing arrangements: opening up tensions between Wikipedia as a carefully constructed encyclopaedia and Wikipedia as a real-time web platform giving up-to-the-minute-information.

03. VISUALISING WIKIPEIDA

As I suggested earlier, the fact that Wikipedia promotes openness and makes these reams of data available to view presents both an opportunity and a challenge to researchers. It is simply not feasible nor productive to read through 1000s of versions of an article, let alone multiple interacting articles. Researchers must find ways of identifying points of interest: what is relevant to the unfolding of the controversy. But as I argued earlier, what is most relevant to the controversy is not necessarily the most easily analysable with respect to the platform. Also, Wikipedia's available information is split between three tabs: the article itself, the history page and the talk page all of which are formatted differently and offer different possibilities for research but are hard to stitch together.

As Wikipedia is known as a site for the staging of controversies, there is a growing field of empirical study devoted to detecting controversy on Wikipedia through automated tools. Yasseri and his co-authors (Yasseri et al., 2012) for example associate controversy with the 'burstiness' of activity in the edit page. In another paper (Sumi et al., 2011) the authors develop an algorithm based on edit count, revert or mutual reverts (when editors undo each other's changes – AKA an 'edit war'). Laniado et al (Laniado et al., 2011) instead locate a certain character of activity in the structure of talk page conversations – that is the depth of indented replies. Borra and Weltrevrede (Borra et al., 2014) have synthesised many of these findings to locate controversies over particular objects (specifically internal links) in an article. Like many digital tools, most of these visualisations are produced as an end goal or as a way of communicating findings to academics, or in the case of Contropedia, with Wikipedia editors and readers. However they are also very valuable for qualitative researchers or virtual ethnographers looking to delineate the study – locating various actors and moments to focus on. Yet the argument of this chapter is that if researchers are to maintain a

tension between studying platforms and controversies then it may be necessary to diverge from the most readily available objects offered up by the platform and cut across different data formats. For example, basing controversiality on frequencies of edit counts (described in the previous chapter as 'liveness') is an easy measure to start with, one which may register certain aspects of the controversy, while mining the less formatted text of the talk page, in terms of depth of indentations, requires relatively more data cleaning and analysis work but also may yield a more nuanced understanding. However, with the exception of Contrapedia, which reveals relationships between the frontstage article and the backstage edit history page, few of these studies question the data boundaries between Article, Edit History and Talk Page. They analyse one area but not the others.

When one develops metrics linked to platform, these may presume to some extent that a device is functioning according to its standard operating procedues. Yet controversies will play out differently in different situations, depending on interactions with other platforms and the specificity of the controversy itself. So, in the spirit of orienting the study to controversy, I will engage with some of these metrics but for now keep open the question of which data formats or sections of Wikipedia are most pertinent to the particular case.

In order to locate a starting point, to identify places where the controversy is 'hottest' on Wikipedia it is necessary to start with a rather blunt measure of frequency of activity, which will be questioned as I proceed. As mentioned earlier, one of the ways Wikipedia editors themselves profile articles and locate controversy is through edit counts and size.[112] These can be visualised with Wikipedia's built in tools such as Wiki Page Statistics, but for more fine-grained analysis I will use the DMI's Wikipedia Scraper Localizer Tool which downloads the entire edit history page of an article as a .csv spreadsheet.

For the page 'Fukushima Daiici Nuclear Disaster', the following graph shows the cumulative edits to the article (red) and the size of the article (blue) in bytes over the course of a year.

[112] These are like 'value meters' in Latour and Lepinay's analysis of Tarde's economics – in that they do not very effectively measure what they intend to but are so pervasive that they effect or format interactions. The measure of article size dominates many Wikipedia discussions.

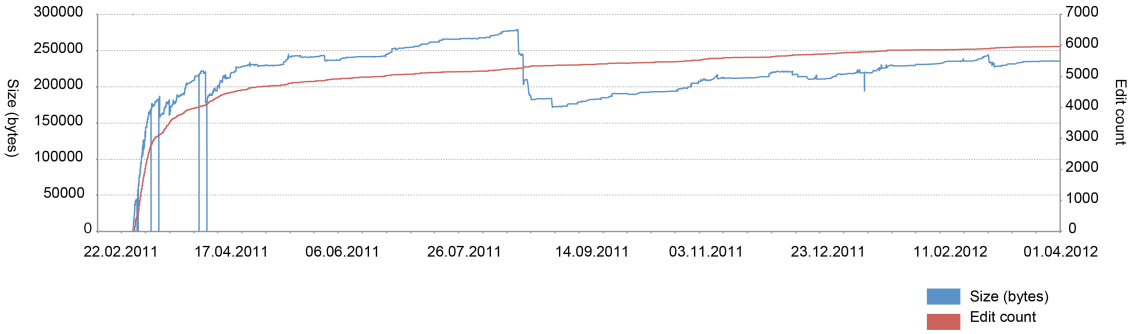


Figure 3. Edit Counts v. Size https://goo.gl/FHS6w3: Cumulative edits and size (in bytes) of Fukushima Disaster Article

Firstly what this simple graph (Figure 3) shows is that much of the article was written while the event was unfolding. There is a rapid pace of editing and expansion in size in the first week, both of which plateau in the ensuing months and years. This is in contrast to the normal, methodical pace of constructing articles about settled historical events.

But what is particularly interesting about the article size are the noticeable dips to zero bytes (the vertical lines). These show when the article was temporarily deleted and the contents merged back into a different article 'Fukushima Daiichi Nuclear Power Plant' where the discussion started as a subheading before spilling out into its own page. This draws attention to another way that this controversy destabilizes the analysis: it shows how the action is not contained by the page, the normal unit of analysis for most Wikipedia scholars, but may prompt the researcher to look at multiple pages.[113] It also may require looking back in time to find the epicenter of the controversy. Finally, this also makes clear that whether or not Fukushima was controversial at all, is not only a question for researchers but part of the controversy itself – is it 'notable' enough to warrant its own page?

But activity measures such as a spike in the frequency of edits (or a sheer drop) could mean many things. It could on one hand indicate the controversiality of the offline activity over Fukushima or they could indicate some Wikipedia effects, such as a petty interpersonal dispute between editors, which may be less relevant to the controversy. To understand what dynamics lie beyond these frequency measures, I will turn to an aspect of Wikipedia which is less easily calculable: the text accompanying the edits.

[113] The Digital Methods Initiative has in fact created a tool which visualizes precisely the relational character of page-disputes (Currie, 2012).

Edit History Page

The Edit History page logs every edit to an article with a time stamp and brief comments, which can help us in understanding what the flurry of editing activity is about: is it regarding the substantive controversy or more banal, procedural elements? On the original 'Fukushima Daiichi Nuclear Power Plant' page, where the event was first mentioned: a lone anonymous editor adds the following sentence to the pre-existing article:

> NHK broadcasing reports: Nuclear Emergency was declared reactors were all shut down but reactor 1 had cooling problems 'cooling pumps' were non-functional

Notice the very un-encyclopedic lead in 'NHK reports'. NHK is a media source but a 'reliable' one, the equivalent of the public service BBC in Japan. A half hour later another editor changes the phrase 'no immediate radiation leak was reported.' (which of course leaves room for radiation leaks happening but not being announced) and adds the cautionary sentence: 'However, there is no evidence of any radiation being released' and references it to a story on *Business Insider*. Other users start referencing the announcement of the evacuation zone but there is a confusion over the numbers of evacuees 2,800 or 5,800 because different sources quote different figures.

The editing starts to accelerate as the possibility of a radiation leak grows. Certain users seem to be itching to report a release of radiation while others like the Administrator Edison are more cautious and constantly remove claims that may be jumping the gun. Administrators are editors with special privileges, such as the ability to ban users who misbehave. Edison's changes are generally respected because he (or she) is a very experienced editor.[114]

As the ordained 'experts' arrive on the scene in the news accounts, there is another edit war over the following statement:

> 6:29am Saturday (JST), local time, anti-nuclear expert Kevin Kamp explains the nightmare scenario in Fukushma, via Forbes and the Institute for Public Accuracy: 'The electrical grid is down. … Given the large quantity of irradiated nuclear fuel in the pool,

[114] Edison must have been around from the early days of Wikipedia to secure the coveted user name 'Edison'.

> the radioactivity release could be worse than the Chernobyl nuclear reactor catastrophe of 25 years ago.' [13]

This quote again refers to the *Business Insider* article which in turn references *Forbes* magazine where Kevin Kamp is listed as affiliated with the organization Beyond Nuclear, an anti-nuclear group. Edison again:

> removed addition literally copy pasted (with a single quotation mark); by a anti-nuclear expert (according to the source); and a worst case scenario (let's stick to what actually is happening for now..)
>
> Kevin Kamps comments are speculative and biased. A life long anti-nuclear campaigner is not an ideal source.)

It is a typical move in potentially controversial Wikipedia articles to do a background check of the sources of quotes to see if they have explicit positions on a topic or interests. In this particular case, the orientation of the expert disqualifies him as a source, despite *Forbes*, a generally reliable source, having already vetted him.

In the first few hours of the nascent controversy, two things become clear. Already there are two sides to the debate, editors who are playing up the severity of the crisis and those wanting to play it down – so as not to start a panic – much like the bureaucrats in Farias' account of the Tsunami Response. Secondly, it is interesting to note how dependent these editors are on the mainstream news and how much of the activity, cosmetic edits aside, is structured around incoming references.[115]

In the spirit of ANT analysis, one can also trace the chains of information flows 'upstream' back to their source. The *Business Insider* article, which is used to temper the original claims about radiation and back up several further statements, is interesting in itself. Not only is *Business Insider* an odd choice of publication to cover this particular event, it is also in fact a 'live blog' – a common format in online news sites. When an event is transpiring in real time, rather than craft separate finished articles, the same permalink can be updated with paragraph-sized tidbits in reverse chronological order (newest on top).

---

[115] Something which is confirmed by a quantitative content analysis of the discussion page (Hara and Doney, 2015)

Scrolling all the way down, it emerges that the first claim about the incident is sourced by *Business Insider* to *Kyodo News* with a hyperlink. The second claim about there not being any evidence of radiation leakage is in fact sourced to Twitter, to the Sky News Account who is in turn quoting a government official.

> Japan government official says technicians are currently unable to pump water to cool the reactor at a nuclear power plant in the country.
> **Sky News Newsdesk**
> **Verified account**
> @SkyNewsBreak

So tracing the chain from Wikipedia through different platforms and sources, the trail leads to at a material press conference in Japan. The irony is that despite the plethora of different sources in the early stages of the Wikipedia article, most of them are in fact referring to *the same source*, the government, who in turn is relying on energy company TEPCO for the latest information.

One can already see that the normal gradations from upstream to downstream information are becoming confused. Firstly a *journalist*, who happens to be mimicking a *blog*-style report is relying on claims made in lowly *social media* about an official *press conference* in which the *scientific content* is already packaged for the public! It is also becoming clear that in the initial controversy over the severity and implications of the leak, participants seem focused around the selection and processing of external *references.* This point of focus is not always the case for Wikipedia controversies but may be more often the case in breaking news stories. Wikipedia allows for the possibility of tracing these information flows back to their source, but this is a manual and arduous task. As I will show next, the *automated* analysis of external references is not easy and researchers who limit themselves to quantitative analyses might then confine themselves to the Wikipedia domain, when the terms of the debate may be already set from outside by online news sites: just as some of the key battles in the settlement of scientific facts lie outside the laboratory in the so called legal or political sphere.

Wikipedia's References

So what is the best way to understand the process of source selection? Kildall and Stern (2011) draw on Latour's analysis of the construction of scientific fact in *Science in Action* (1987) to analyse how Wikipedians use references to secure facts. According to this model, as facts become more accepted, they will require less references and attributions to prop them up. This means that a citation might be qualified as 'Source A claims B' then become 'it is generally accepted that B' then a simple reference followed by a unqualified statement of fact – which is to say that less references are required as controversies cool down. This is an interesting proposition and one which would be generally consistent with Wikipedia's self presentation as 'constantly improving', but as I pointed out earlier, the job of source selection has as much to do with practices in journalism.

For example, in the seminal study *Policing the Crisis*, (Hall et al., 1978) the authors explain in relation to the issue of mugging that the news favours 'elite' versions of events because the professional need for 'reliable' sources made the news dependent on official spokespeople and accredited experts. Professional standards did not make things 'objective' or 'neutral', they help news media conform to the dominant view. In Hall and his co-author's terminology, these official spokespeople are often 'primary definers' of the topic because any alternative, or indeed all subsequent, accounts needed to either inhabit or acknowledge their framing of the topic or be relegated to the status of a marginal or alternative view. Now ANT researchers would not assign dominant class position or interests to particular sources, but it does hold that the first actors on the scene have a privileged role in defining a 'problem' (Callon, 1980) or issue (Marres, 2012a) and that certain types of sources may be favoured through routines, policies and technologies, though this is always negotiable.

However there is not an existing tool or method for visualising Wikipedia's references. They are 'digital traces' but not 'digital data' in Venturini's terminology. The referencing system is fully integrated into the interface – users enter it with a <ref> tag in the text and they are automatically collected at the bottom, but there is not a universally recognised format for the citations:

> <ref name='reuters'/> some 300,000 people [[Emergency evacuation|evacuated]] the area; 15,884 (as of 10 February 2014)<ref name='cnn140220'/>

Some references include a published date but not an accessed date; some do not have authors or even hyperlinks. This is exacerbated in fast paced events where some users may place a bare hyperlink in the hopes that another user will clean up the reference later. Despite this, the references are an important influence on the content of the article and which may reveal patterns over time, so it seems pertinent to try and visualise them – to go somewhat against the formats supplied by the device.

The below visualisations were created by scraping different versions of the two articles I have looked at so far: 'Fukushima Daiichi Nuclear Disaster' and 'Fukushima Daiichi Nuclear Power Plant' for the URLs contained in references at the bottom. Due to file-size constraints, I was only able to scrape 1 in 10 version of each article.[116] For each version, I collected each of the references and parsed the host domain from each full URL (e.g. – 'bbc.co.uk' from 'http://www.bbc.co.uk/news/world-asia-pacific-13678627'). References, which did not contain a link or where the scraper failed to obtain one, were labelled 'No Link'. I then visualised the composition of sources, by which I mean host domains, for each version using a 'stream graph' provided by Density Design's RAW application. I began with the first 4 days of the article where activity was concentrated according to the frequency diagrams. This would contain the major events, such as the explosions and meltdowns and would focus on the primary definers.

[116] This is a practical limitation. It would be preferable to use continuous data, or at least group and sub total the references from every 10 edits. Sampling 1 in 10 allows for the unlikely possibility that something catastrophic is happening between them that will not become visible.

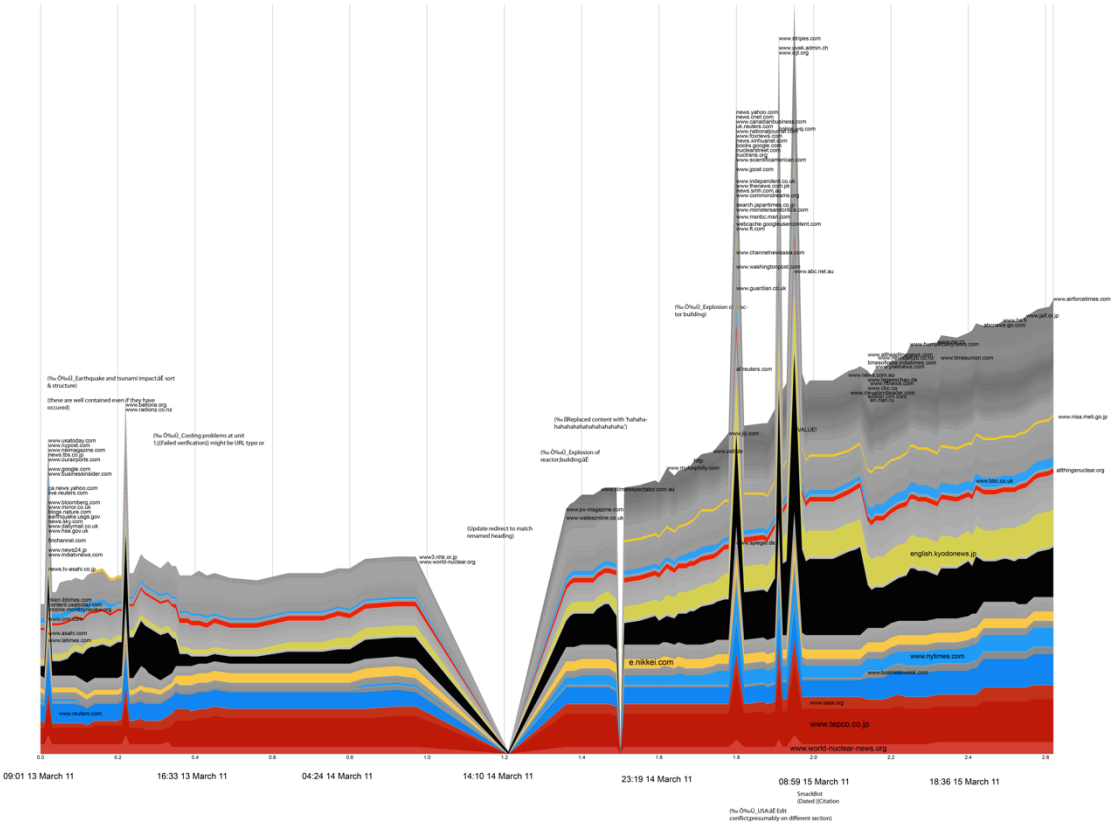


Figure 4. Fukushima Daiichi Nuclear Disaster References First Week

https://goo.gl/f0E1Ar: Composition of article references March 11 -18

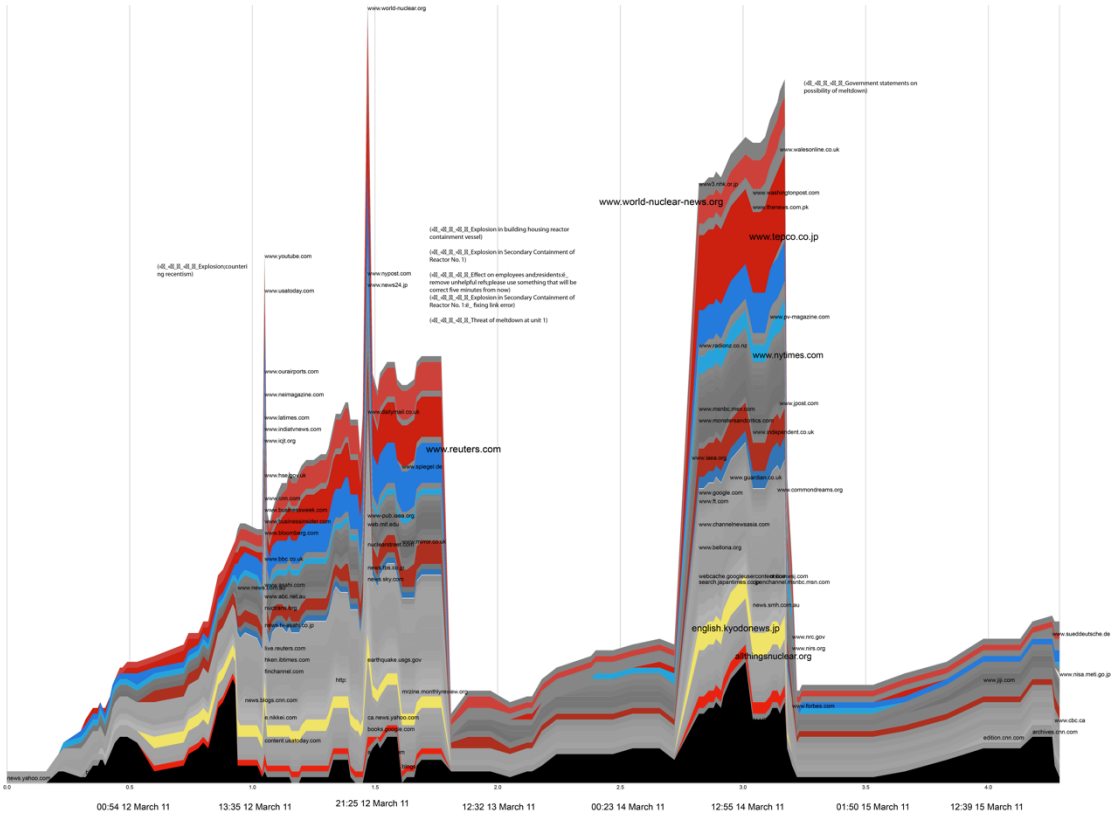


Figure 5. Fukushiima Daiichi Nuclear Power Plant References First Week

https://goo.gl/iFSXC6: Composition of article references March 11-18.

In the above images, the x-axis represents time and each vertical slice represents an individual version of the article. Each domain (e.g. – www.bbc.com) is given a coloured strip, sized according to the number of individual URLs from that domain contained in the references.[117] Following RAW's conventions, the stripes in the top image are ordered from largest to smallest, largest on the bottom, starting from their position in the first slice, while in the second image they are arranged in reverse alphabetical order. Select domains were hand-coded to highlight different types of sources: red = nuclear industry sources, yellow = Japanese news sources, blue = Western news sources and green = environmental organisations. Keep in mind that the top graph 'Fukushima Daiichi Nuclear Disaster' and bottom graph 'Fukushima Daiichi Nuclear Power Plant' are in a slightly different time scale as activity actually starts a couple days earlier on the 'Power Plant' page; and also a different scale in terms of references – the 'Disaster' page has almost three times as much.[118] The black strip going through both graphs represents references with no hyperlinks (either references to physical books, or perhaps, errors in the scraper). Keeping these accidents of the process in view reminds the viewer of the conditions of the graph's production.

This graph does manage to flesh out and qualify the frequency graphs, revealing how the composition of references is potentially skewed. Normally, Wikipedia articles are based on books and scientific articles but this graph clearly shows a focus on mainstream, Western news organisations (blue) particularly Reuters. This is not uncommon in Wikipedia articles about breaking news stories, but what is perhaps more striking in this case is the amount of reliance on self reporting by TEPCO, the owner of the reactor, itself in the form of press releases (red). This perhaps speaks to the lack of independent information on the ground. Also playing a significant role was the IAEA (an international UN based agency for nuclear safety). Even for a controversial article, both articles have an enormous number of references 300-400 for the 'Disaster' page, but the quantity of sources masks a homogeneity of voices – in this case mostly official government spokespeople and private corporations.

[117] This graph format is partially inspired by IBM's History Flow visualization (Viégas et al., 2004) which uses coloured strips to indicate the authorship of revisions. This tool however does not focus on the references.

[118] One can also see the severe drops in the number of sources as text and references were moved between the two articles. Because the graphs are spaced according to edit counts not according to time, these currently do not line up.

While a gradual increase in sources over time is to be expected in an unfolding event, what is more puzzling are the rapid spikes in which sources are added then removed. Starting with the bottom graph, the 'Power Plant' page, on the 13th March there is an odd spike in the number of sources, followed by a more gradual accumulation. Looking at the comments on the edit page, it appears that this corresponds with the announcement of an explosion in Reactor 1 which is followed by a jump in the sources coming from TEPCO particularly. Yet the TEPCO sources themselves are not particularly enlightening. Some are radiation readings in tables which are presented without explanation or context (Slater et al., 2012). But this means that the editors are starting to use *primary sources* – that is bypassing the news and travelling 'upstream' to the expert announcements.

In fact nearly all of the sharp jolts in both graphs correspond to particular events or announcements in the news – sources and information are first gathered and then reduced as the significance of the event becomes clear. But what about the drops? The massive gap in the middle of the top graph, the 'Disaster' page, represents one of the attempts by editors to move the page back to the 'Power Plant' page – this gap is mirrored in the bottom graph where a host of new sources are added from the disaster page, but later taken away.

It could be suggested that a sharp increase or decrease in sources, is an indicator of controversy, but are these necessarily controversies about Fukushima? For example, around the 15th of March, there is a sheer drop in sources but this time it is not driven by an event but by a bot called DumZiBoT who is in charge of cleaning up references – converting bare hyperlinks (a placeholder, normally) into full citations. The bot also merges duplicate references. This is why the 'no link' bar contracts noticibly because these represent some *broken references* which have been fixed.

> (reflinks: Bot: Converting bare references, using ref names to avoid duplicates, see FAQ) ref names to avoid duplicates, see FAQ

Because this graph relies on frequencies of sources it is not clear if dynamics are due to the controversy, such as the explosion, or due to the platform, such as the automated formatting of links or squabbles over the 'notability' of the event. Some of these

platform controversies are important for the representation of the controversy while others are not.[119]

Next I will examine the main 'Disaster' article from a wider vantage. The below graph (Figure 4) utilises the same approach but each slice corresponds to 1 in every 100 edits and spans the first year of the article.

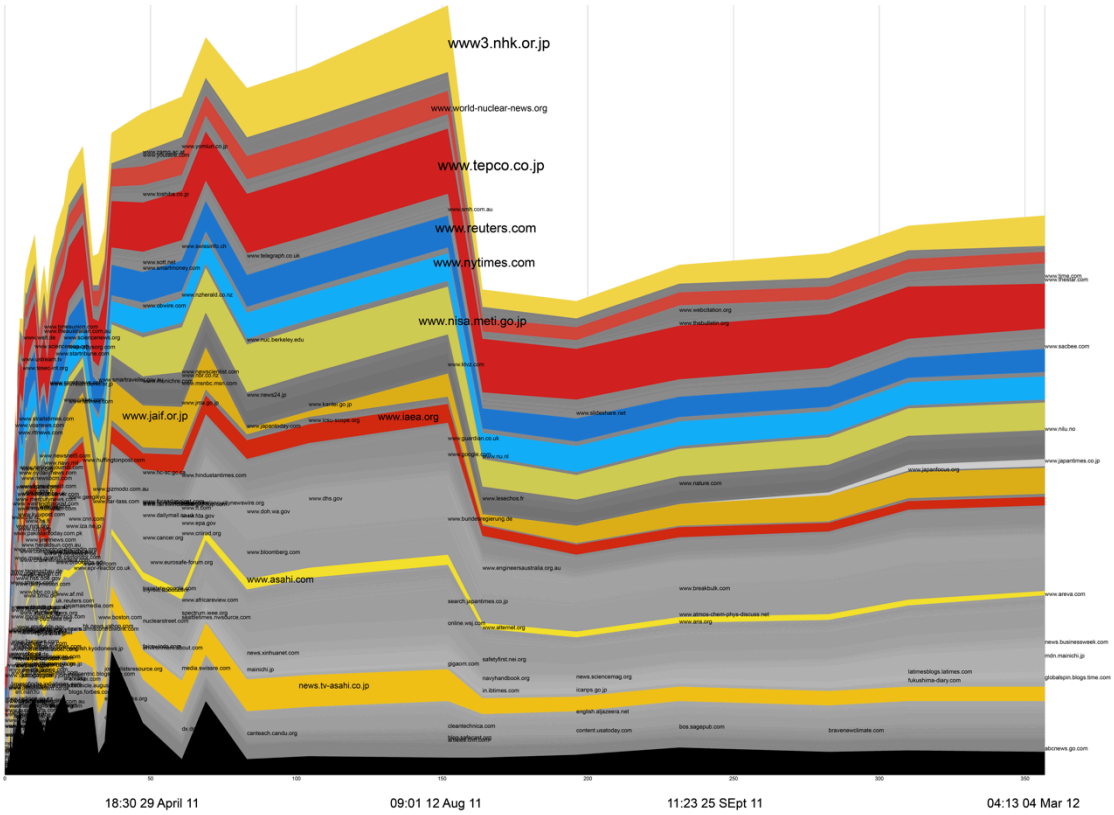


Figure 6. Fukushima Disaster References First Year https://goo.gl/zJNsp2: Composition of article references March 2011 – March 2012 each slice every 100 edits.

The above graph gives a longer but less granular view of the dynamics of references. The time span of the previous graphs is here represented in the flurry of activity at the left, where the time slices are denser – the pace of editing slows considerably as the months go on. Now if Wikipedia references functioned like scientific citations, as Kildall and Stern have suggested, then there should be a process of black boxing, that is the gradual *removal* of sources as facts become broadly accepted. The primary documents (press releases) and wire services should be replaced by the ostensibly more detached mainstream news, which would in turn be replaced by more comprehensive scientific

[119] The second drop to zero in the top graph for example is a case of simple vandalism, where the article was deleted: the accompanying comment reads "Replaced content with 'hahahahahahahahahahahahahaha"

and historical papers and books. While around August 2011, there is a systematic culling of sources and cleaning up of the article, which was also visible in the frequency graphs, what is interesting is that they have not been replaced but merely built around. The original Reuters and TEPCO stripes remain largely intact as other less prominent sources were pruned around them.

This approach at least seems to suggest that the 'primary definers', which largely originate from official or government sources, *were* in a position to set the tone for the ensuing discussion with their central placement among the sources. But this visualisation cannot, on its own, explain the impacts of this on the content of the article or why for instance, blogs or alternative voices are relatively marginal while insider primary sources are so prevalent? If the sources often come from TEPCO, then how does the article maintain such a critical stance on the company? Next I'll look at how individual sources are parsed by the Wikipedia community. This requires analysing yet another area of Wikipedia which is less readily analysable, even though it is technically open for all to see.

04. JUSTIFYING THE SOURCES: TALK PAGE

While it is important to note these asymmetries in Wikipedia's references from the very early stages of the article, this must be supplemented with an understanding of the process through which sources are selected. For this I will now turn to the other tab on Wikipedia, the 'Talk' page, which is where problems, which could not be resolved in the edit page, are discussed in more detail. The discussion section is arranged in headings with posts under them, which are time stamped and 'signed' by a user or an anonymous account (identified by an IP address). Replies are denoted by an indentation under a comment. The talk page, like the references is a textual artefact, which is not as easily scraped. Laniado and his co-authors, mentioned earlier, analysed the talk page in terms of the depth of indentations – the number of nested replies – as a way of locating controversy. This is a helpful visual guide when scanning the page, but in this section I will resign myself to qualitative textual analysis. Due to the formatting differences between these areas of Wikipedia, it is difficult to link up particular article versions or particular edits to discussions on the Talk page and this requires patience and methodical qualitative tracing.

I thought of this analysis in in relation to White's 'gatekeeper' (1951) study. White asked a wire editor, sifting through a feed of possible stories, what his reasons were for accepting or rejecting them. Most of his decisions were practical or stylistic but some reflected 'ideological' or cultural 'biases' – the latter of which became overstated in the scholarship that followed (Reese and Ballinger, 2001). So in the case of collectively 'gatewatching' (Bruns 2005) with online media, what justifications or rhetorical strategies do Wikipedia editors give for erecting boundaries between fact and fiction or reliable from unreliable and, diverging from White's study, what non-human technologies participate in this process?

As I will show in this section, the Wikipedia editors exercised some autonomy in their source selection, negotiated within parameters set by both policy requirements and the technical affordances of certain webpages. I will focus, again on the very beginning of the discussion to understand how the editors negotiate the uncertainty and because this segment of time contains many of the types of strategies present in the later stages. I will quote these discussions at length because the editors are actually quite articulate at explaining their reasons for selecting sources over others.

News Source or Encyclopaedia?

One criteria for source selection hinges on the tension noted earlier between Wikipedia as an encyclopaedia and as a news medium. Starting on the 11 March, with the branching of the article from 'Fukushima Daiichi Nuclear Plant', the 'Disaster' article is quickly beset by complaints about Wikipedia's role in relation to the news. This mainly manifests itself as a problem of using present tense, no doubt a symptom of editors paraphrasing from the news, while encyclopaedias will conventionally speak of events in the past tense.

One editor at the end of 11 March (UTC) twice removes a direct quote from a source speculating about the consequences of the disaster, arguing that:

> we should focus here on the facts and immediate risks rather than discuss risks if events don't unfold positively in the next days in worst case (the name given in the article it cites....). Views welcome!
>
> L.tak (talk) 22:03, 11 March 2011 (UTC)

Wikipedia, as an encyclopaedia is inherently focused on the past but the controversy over the effects of the disaster has awkwardly turned the community towards speculating in a more journalistic mode. The temporal range of Wikipedia's articulation of Fukushima has implications for what possible *long term* effects can be claimed, as in the case of the WNN article discussed earlier.

Soon after, the news / encyclopaedia uncertainty is broached more directly:

> The article is full of the latest information, appropriate for keeping on top of things. …this site is supposed to be an encyclopedia, which is appropriate for getting to the bottom of things.. —Preceding unsigned comment added by 71.59.236.139 (talk) 05:42, 12 March 2011 (UTC)

But a few days later, another user in the same discussion thread defends what Wikipedia is doing:

> This is the only location on the internet I'm aware of that consolidates the facts of this ongoing event in a concise and complete way. News stories from the standard sources are actually a poor way to follow what is happening, because each story is 95% the same content as the previous, with only a few new facts added as events unfold. At this moment this article is on the front page of Google News, and that is out of 16,853 other articles concerning this nuclear accident. So this article is being read by many, many people, and even Google recognizes the value of the information covered here. --Dan East (talk) 01:06, 15 March 2011 (UTC)

This quote is interesting, firstly because the user defending the article argues that Wikipedia's value added is the *consolidation* of information which is otherwise duplicated and rehashed. Interestingly, Google rankings are deployed as a measure of the importance / audience size to justify continued effort in improving things. As discussed in Chapter II, invisible or lurking audiences (who do not actively participate) are an elusive entity (Ang, 1992) who are nonetheless frequently invoked to settle arguments, this time materialised through Google Rankings.

Industry / Government Sources

So how did the editors justify the use of primary, industry sources? Partially this has to do with a perceived lack on the part of mainstream news:

> Presumably the mainstream media will improve their coverage in the next day or two as they come to better understand the technical issues, but for now, this is the most reliable article I've found. It appears to be updated every several hours.
> --A. B. (talk • contribs) 23:57, 11 March 2011 (UTC)

Because the editors do not trust the mainstream media's ability to process the information they first consider reports from the WNN but are concerned that it's specialist language will be too rarefied for their readers. The editors are making judgements about appropriate language, in the same way as scientists make judgements about 'appropriate simplifications' in media (Hilgartner, 1990).
The same editor then offers the TEPCO press releases, which he correctly notes are the source for most mainstream news anyway. Administrator Edison, who we encountered earlier, is suspicious of TEPCO's use of the term 'elevated radiation levels ' but then resigns himself (or herself) to the fact that they must rely on them:

> We should restrain anyone's going beyond official statements and reliable press coverage, in the article text. Edison (talk) 00:51, 12 March 2011 (UTC)

Much like journalists, Wikipedia editors are forced to rely on 'official sources' but they are doing so ambivalently and strategically, like Brian Wynne's sheep farmers cooperating with the scientific experts (Wynne, 1992).

Expertise

Perhaps the most common way of refuting sources in this article stands in sharp contrast to the account of NPOV and reliable sources given in the policy pages: through the expertise of the editors.

> There is another error: 'electric power for the cooling turbine' is clearly wrong as the turbine is the passive element and does not generate heat. It is the fuel rods or elements (mostly uranium oxide packed in zirconium capsules that are packed into hollow tubes that are mechanically adjusted to get criticality) that need cooling...

This editor is contesting the article's wording, not based on another source, but on his implied expertise. He is presenting a version of the world which the article is at odds with, ushered in on the back of a authoritative tone. Because of the anonymity of most editors, it is unclear who is an expert or not, so this must be rhetorically performed. Speculation without reliable sources is supposedly forbidden in both articles and discussion pages but editors often use their own outside knowledge to refute the accounts of even respected news organisations. But, sometimes this goes beyond critiquing representations of journalists to questioning the interpretations of the scientists and nuclear experts themselves. In a post on the 15th of March, the editors bypassed the media and even the government spokespeople, who they are convinced are wrong, and started to create their own graphical representations (Image 2) of the numbers, direct from Tesco's press releases.

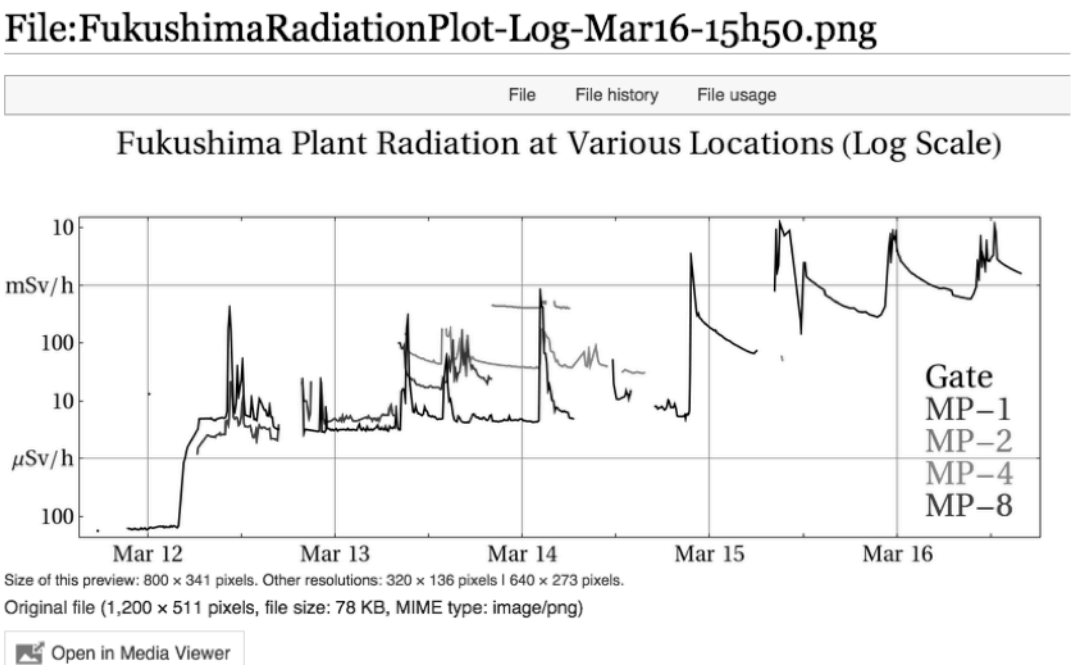


Image 2. One user's graph of radiation measurements, coded by measurement device.

But rather than packaging or simplifying the figures for the audience they present the 'raw' data from the monitoring posts themselves, along with their geographic location, opening up the monitoring infrastructure to outside scrutiny, much like the Japanese citizens did in Morita, Blok and Kimura's account (2013).[120] Were the experts and journalists cannot be trusted, the editors use the principle of openness to collapse the boundaries separating scientists from journalists and audiences. This is again helped

[120] Of course there is no such thing as 'raw data' and the placement of the monitoring posts and the choice of representation already structure possible interpretations and responses.

by the fact that it is not *a priori* clear who is an expert or not. So even though expertise, in the sense of institutional credentials and official spokespeople, has been problematized on Wikipedia, expertise can be mobilised rhetorically to critique the institutional experts.

Accounting Practices

However, while expertise has been problematized, Wikipedia has other socio-technical boundaries, such as between backstage discussions on the talk page and the audience-friendly article itself. While the 'Talk' page is full of transgressions between news and scientific norms and expert and journalistic roles, these messy exchanges, much like scientific 'shop talk' must be somewhat sanitized before crossing over to the 'Edit History' page and affecting the article itself, under the watchful eye of bots and other monitoring devices. While the editors may select references based on expertise, they must *account* for their choices through conventions, policy and technologies which sometimes re-assert boundaries between different types of media.

This most obviously involves the policy of Reliable Sources, discussed earlier, and the preference for books and scientific journals over blogs, but also this involves the policy of Verifiability. Sources must be in English so they can be confirmed by the average editor, which in this case necessarily excludes some local Japanese news sources, but not, interestingly, TEPCO's press releases. So these conventions are somewhat malleable, at least in unfolding events.

However, some of the most consequential reasons for citing or challenging a source have to do with their technical features.

> I removed a reference to the BBC 'live blog', since I think it's unverifiable. I notive [sic] 5 more references to it. Thoughts? 220.100.15.15 (talk) 02:47, 13 March 2011 (UTC)
>
> BBC is always reliable —Preceding unsigned comment added by 24.18.132.80 (talk) 04:33, 13 March 2011 (UTC)
> Unverifiable is different from unreliable. That link is unverifiable, because it is dynamically updated. 113.197.242.129 (talk) 05:25, 13 March 2011 (UTC)

Despite the fact that the BBC are one of the most vetted sources, as the IP address alludes, live blogs are frowned on because their content is dynamically updated and

may not always contain the referenced information — or that information will become hard to find, buried in new material.

This means that the technical affordances of Wikipedia and other platforms impact source selection. For example, websites that use permalinks for articles are always preferred to sites where the content is liable to change. There are also numerous bots, scripts, templates and tools for formatting and locating sources, some of which may incline editors toward certain common source types.

One of the most important tools for sources is perhaps Google.

- Editors occasionally use sources with the links hosted on Google News:
- When the TEPCO website was temporarily down (possibly due to the extra traffic) Google was used to access cached copies of the press releases.
- When editors are expecting or anticipating certain sorts of coverage, they will use Google to find a news source meeting their criteria (e.g. – Fukushima + Meltdown)
- Google results are even used to decide on 'notability' of statements or reliability of sources.

So in many ways, Wikipedia legitimacy is fed back through Google's algorithmic legitimacy (Gillespie, 2014). This could be one possible explanation for why 'official' or 'corporate' sources are preferred: because they likely have better technical infrastructure which might be favoured by Wikipedia's policies or Google's indexing, as opposed to home made blogs or independent sites.

To only look at cultural explanations of Wikipedia's process, would be to miss the contributions of technical features of the platforms and also miss the contingency of particular controversies. Overall, policy and technical features limit, though accounting practices the overall spectrum of sources to the official, largely mainstream media account of the event – reinforcing the upstream-downstream model of science communications. However, as I showed earlier the editors can in practice shuttle all the way upstream and back downstream as long as the boundaries of the dominant view are put back in place later.

05. CONCLUSION

This chapter was concerned with an empirical case in which socio-technical boundaries between science and media and between journalists and non-journalists were destabilised and renegotiated through acts of boundary work. Yet in order to understand this complex situation I argued that it was necessary to question another set of socio-technically constructed boundaries in the data structures of the platform itself. ANT-informed approaches insist on a certain freedom of movement on the part of the researcher. In some respects Wikipedia with its wealth of time-stamped and formatted data makes it possible for researchers to trace links from claims made in an article back to their conditions of production and the original sources on other websites, particularly online news and even back to press conferences and the testimony of experts. But in another sense Wikipedia directs this analysis by making certain data more or less analysable from either a quantitative or qualitative perspective. Ironically, the programme of openness or transparency of information on Wikipedia which enables this form of research, actually results in divisions and silences in the information presented. Edit counts can be graphed with the push of a button while less structured references and largely unformatted text are less amenable to being visualised. Also because that the article, edit history and talk page are not directly link together or 'joined up': it becomes difficult to trace controversies across them.

The advantage of analysing online platforms in relation to controversies is that they not only allow us to study how online platforms possibly intervene in controversies, in this case by selectively highlighting information and contesting expert accounts, controversies may also to some extent destabilise the device, making previously black boxed technologies and practices available for analysis. For example, it was only through the specificity of the controversy that Wikipedia's complex system of vetting reliable sources came into focus. This led me to study the larger ensemble of news, blogs and press releases which are not easily analysable with digital tools. For this reason I created a data visualisation which visualised references to external websites. This was important because, in this particular event, Wikipedia was dependent on various online news sources as reliable sources in the unfolding Fukushima disaster, particularly sources originating from within the Japanese government and nuclear industry, which is against Wikipedia's normal protocol. Yet the lens of particular platforms and particular devices such as references also teaches us about

controversies, in contrast to past nuclear incidents, the controversy was not settled by accredited experts or facts but through negotiations over reliability in relation to the news.

By following controversies over source selection across Wikipedia's three separate tabs (article, edit history and talk), I was able to show that although boundaries between scientific experts and journalists often become entrenched in disasters due to the scarcity of information, in this case at least, they were effectively undermined. Yet boundaries become reasserted in other guises: between reliable and unreliable sources, experienced and inexperience editors, particularly as a form of accounting practices.

Throughout this chapter I oscillated between the polls of controversy-centred and device-centred analysis. A controversy over radiation, which bubbles up in online news and blogs, manifests itself on Wikipedia as a media-specific controversy over the bias of an anti-nuclear source. Similarly the media-specific controversy about whether Wikipedia is an encyclopedia or a news source affects how Fukushima as a matter of concern is articulated (whether or not they are allowed to speculate about the future in a more journalistic mode). Neither the frequency graphs nor the area graphs of the references however could speak to this tension, which only became apparent through qualitative analysis.

There are, however, several limitations of the admittedly crude graph I produced. Firstly, it was necessarily grouped into slices, that is sampling every 10, or 100 edits, when of course significant changes in the source composition could occur between these slices. Secondly, the fact that I have hand coded the categories, as a necessary reading aide, tends to reify categories like mainstream media and blogs when the boundaries between these two are being complicated empirically. It would be preferable to follow categorisations presented by the research subjects or the platform (such as domain type .com .gov or perhaps country if possible in the future). Finally I was forced to first map references and *then* qualitatively analyse the discussions around their use separately. Ideally this would happen together tracing individual references to their justifications, but Wikipedia enforces a division between these two forms of data and it is hard to stitch them back together.

As a lone researcher with limited programming ability, I could only resist the formatting of the platform so much. But this has to do with the specific constraints and affordances of Wikipedia. In other chapters I will go further in combining qualitative and quantitative work, create more continuous representations of time and more granular categorizations which avoid nominal categories like 'mainstream news'. In the next chapter I will make the point that while it is important to move past easily quantifiable 'digital data' to take in less formatted data when the controversy demands it, it is also wrong to dismiss these quantifiable digital traces all together because they also act on the situation, something which is increasingly important in so called social media platforms which are increasingly quantified.

V. WHAT 'COUNTS' AS PARTICIPATION: DIGITAL DEMONSTRATIONS ON FACEBOOK

The previous chapter argued that tracing controversies through online platforms might require traversing boundaries which are materially enforced by platform architectures and data formats: especially attending to digital traces which are less accessible and formatted, even going so far as to quantify them. But while it is important to not confine studies only to the 'low hanging fruit' of easily quantifiable metrics and structured digital data, we also ignore these traces at our own risk. It must be acknowledged that these metrics, markers and numbers do not just describe but shape and reflexively mould interactions on these platforms. Just as university rankings reshape the institutions they are meant to monitor (Espeland and Sauder, 2007) online traces such as likes and edit counts may encourage certain sorts of behaviours over others. Confronting the performative effects of data is especially urgent given the increasing quantification of activities in online media, and the pervasive influence of metrics and rankings in these spaces (Gerlitz and Lury, 2014). So in studying online activities, I argue that it will be important to *qualify* these quantitative traces by bringing them into dialogue with other sorts of data. This means, again deploying methods in ways not intended by platforms.

Decentring these devices, as I have proposed, becomes easier when they are analysed in relation to specific contingent controversies. This chapter will focus on the controversy over the proposal of a new nuclear power plant in the UK at Hinkley Point, Somerset: a long running dispute which became reignited in the wake of the events at Fukushima. This elicited regular co-ordinated protests on the ground as well as on-going online activities. While these of course involved sub-controversies over knowledge and scientific *representations*: over the existence of clusters of childhood leukaemia in the proximity of nuclear plants and the presentation of nuclear as a low-carbon technology, they also involved controversies over *representation*, or participation in decision making in lieu of official venues (Johnstone, 2014). So this object straddles conceptions of more science focused knowledge controversies and more public issues and once again may require a broader lexicon for describing it. I will

examine the unfolding of these controversies mainly through Facebook, one of the key social media sites as well as email lists and static webpages, which are just important for these activities. Although I will address the usefulness of these platforms for offline activism, how these platforms are implicated in organizing direct action, I want to focus on the mainly online interventions, which have more ambiguous effects and objectives.

First of all I will again pose the question of what contribution online platforms can make in controversies, especially in relation to activism, starting with a discussion of the pejorative term 'clicktivism' and how online forms of activism present conceptual and methodological challenges. Starting with an analysis of email lists I will then discuss some ways to resolve this using an idea from Warren Sack and some bi-partite networks. Next I will turn to Facebook pages using a similar approach. However, the increasing quantification of Facebook makes it more pressing to resolve these quantitative / qualitative tensions, something I will do with an innovative data visualisation. I use this visualisation to show how quantitative traces increasingly draw activists into a visibility game which they cannot hope to win against larger actors like corporate webpages. As discussed in Chapter III, they may also draw researchers into the study of 'popular' or 'trending' content (Marres and Weltevrede, 2013).

I will not, it should be said, go into very much detail about how these traces shape the calculative capacities of platform users, nor will I be able to discuss fully the affective or symbolic practices which these traces may elude to: the goal of my approach is to develop techniques which leave space for these sorts of phenomena in the analyses.

01. 'CLICTIVISM' AND ACTIVISM ONLINE

In activist circles the potential contribution of online platforms to politics has frequently been derided as 'clicktivism'. 'Clicktivism' was most famously coined by Micah White in 2010 in two articles in *Adbusters*[121] and *The Guardian.*[122] In both articles, White warned that left activism, in embracing online tools of engagement (signing online petitions or just clicking 'like' on a Facebook post) was buying into the logic of marketing.

[121] Available at: https://www.adbusters.org/blogs/blackspot-blog/rejecting-clicktivism.html (accessed 7 September 2015)
[122] Available at: http://www.theguardian.com/commentisfree/2010/aug/12/clicktivism-ruining-leftist-activism (accessed 7 September 2015

> 'This manifests itself in an inordinate faith in the power of metrics to quantify success. Thus, everything digital activists do is meticulously monitored and analyzed. The obsession with tracking clicks turns digital activism into clicktivism.'

This calculative monitoring of participation (White specifically had in mind the site MoveOn.org) is the crux of his critique of online activism, but has been somewhat lost in more famous discussions which contrast online activism with traditional offline activities (Gladwell, 2010; Morozov, 2012). It is worth remembering that White was never against online activism *per se* only its quantification: in the wake of the Arab Spring (Castells, 2013), White and *Adbusters* were instrumental in launching the Occupy Wall Street movement in the United States in which social media played a central role.

This highlights the two empirical problems I want to address in this chapter, firstly it is important to rehabilitate and better understand how predominantly online activities might contribute to controversies and politics more widely and second, I want to address what is at stake in these activities being defined in terms of metrics. In the last chapter, I raised the possibility that platforms like Wikipedia could intervene in online representations of the controversy, but what about new modes of participation and assembly? It is relatively well established that online platforms, can be used by activists to organise protests on the ground, and this is something I will discuss, but the role of online activities in mobilising and defining movements is more ambiguous (Gerbaudo, 2014).

On the web, as discussed earlier, Marres and Rogers (Marres and Rogers, 2000) theorised that hyperlinks as a device could be used to make visible issue networks – defining the field of the controversy. Marres particularly notes the ambivalence and antagonism inherent in these formations. As discussed in Chapter III online platforms may entail even more tenuous modes of association. In the last chapter I dealt with a very fluid constellation of actors on Wikipedia, some of whom exhibited a sustained engagement with the article while others, including bots, appeared only to fix a 'typo'. On social media, users may be passively directed towards a topic at the suggestion of an algorithm, with only a passing attachment the issues: just long enough to click 'like' or make a wry comment, before dispersing again. Sometimes these nameless assemblies may warrant the name 'publics' (Plantin, 2011) but social media, according

to the literature at least, may entail even more ambiguous levels of involvement.

One positive way of defining these ambivalent attachments is advanced in Bennett and Segerberg's *The Logic of Connective Action* (Bennett and Segerberg, 2012). The authors propose that recent protest movements from the Arab Spring to los Indignados to Occupy Wall Street are unique in the history of social movements because instead of being brokered by 'brick-and-mortar organisations' with membership lists promoting a shared identity or a 'collective action frame' – partially agreed objectives for the movement (see Benford and Snow, 2000)[123] – social media such as Twitter allow the participants to share their often contradictory *personal* demands which are then aggregated, ratified or ignored, rather than synthesized collectively through either consensus or another organised process.[124]

They call this logic in which the medium, rather than any particular message, shared identity or goal provides the cohesion 'connective action' in contrast to traditional 'collective action'.[125] The lynchpin of connective action is the participatory sharing of content, which emerges out of online peer production and open source software communities (Beer, 2009; Benkler, 2006; Kelty, 2005; Reagle Jr, 2010). Bennett and Segerberg have recently built on their argument thorough a quantitative analysis of 20 million #Occupy Tweets (Bennett et al., 2014). The authors discover a kind of division of labour between various content sharing practices (linking, retweeting and hashtagging) which are seen as evidence of distributed organisational capacity amongst the cacophony of voices.[126] But, in contrast, one of the article's respondents (Gerbaudo, 2014) questions the relevance of studying these micro-practices on group cohesion, when the group is clearly '…more than the sum of its parts'. He argues that researchers must take into account the group's 'identity and intentionality' and the larger culture that shapes individuals.[127] Although I do not have any particular

[123] Out of these terms, I prefer the use of collective action frames because this, at least in some uses, comes from a social constructivist perspective: frames are strategically engineered. But I prefer to simply locate framing practices in individual utterances rather than assuming they are collectively held or generated.

[124] These sorts of claims may hinge on the distinction between the broader social movements (SM) or perhaps issue-publics and social movement organisations (SMOs) which may be more clearly defined. In this chapter I am concerned with SMOs.

[125] This sentiment is echoed by Gladwell, but in a negative light: 'Where activists were once defined by their causes, they are now defined by their tools.' (Gladwell, 2010)

[126] Connective action is theorized in relation to Twitter but the implication is that it applies to other social media as well.

[127] This amounts to a Web 2.0 rehearsing of one of the classic debates in social movement studies: between 'collective action' (Olson, 1965) and 'resource mobilisation' (McCarthy and

investment in these debates within social movement studies, what these conflicts highlight is that studying online modes of assembly cuts across dichotomies between micro and macro, and I will argue, quantitative and qualitative approaches, and this makes them particularly difficult to study.

Social Movements as Relational

In relation to the study of *offline* social movements, Israel Rodriguez-Giralt, (Rodríguez-Giralt, 2011) points out that debates in social movement theory are beset by many dichotomies like the above: individual traits versus social tensions; subjectivities versus objective conditions which cause unrest; symbolic versus practical deployment of resources; micro versus macro level of action etc.[128] To be fair to the very complex social movement literature, these scholars are constantly trying to resolve tensions like these[129] but rather than further specifying and qualifying these divisions, Rodriguez-Giralt proposes to cut across them using insights from Actor-Network Theory, which I will briefly reiterate.

Instead of starting with pre-existing groups, one must explain their precarious stability as shifting associations between various actors conceptualized as a continually emergent actor-network (which includes multiple groups or networks). These networks are constantly attempting to enroll new actors and strategically reframe or 'translate' the roles of other actors on their terms. These networks of course include non-humans who may be just as consequential for the definition of the assemblage. In Rodriguez-Giralt's example, the movement gathering around a chemical spill in Spain only achieved international attention when a flock of migratory birds physically transported the toxic chemicals to different locations in Europe – creating and connecting far-flung groups.[130] So when viewing groups, we need to view them in relation to particular controversies and other groups in the issue network.

Zald, 1977): roughly, are social movements better analysed as a matter of organisation and infrastructure (located in micro-practices) or through an interest in the identities or ideology?

[128] There are also divisions between American and European theories (resource mobilization vs new social movement theory)

[129] Rodriguez-Giralt discusses (Melucci, 1996) who similarly argue against some of these entrenched dichotomies but according to Rodriguez-Giralt ends up reproducing them himself.

[130] Isaac Marrero-Guillamon (2013), who also applies ANT to social movements makes a similar point using the work of Gabriel Tarde, discussed earlier. In this formulation, identities are always *relative* to the network and are a consequence of shifting associations rather than a cause of actions.

However, it is important to add a careful qualification when access to groups (and controversies) is mediated through platforms. Following Latour (2005)'s dictum that there are 'no groups only group formation': rather than taking their existence for granted, I hope to investigate the *performative* definition and maintenance of groups online. So this involves searching for 'spokespeople' who have an interest in defining the group (and oppositional anti-groups) and following the '…traces left behind by their activity of forming and dismantling groups' (Latour, 2005: 29). Of course one of the main spokespeople for groups are the social media platforms themselves, which articulate these entities through pages, and importantly, metrics, which I will come to later.

To help illustrate this, I will discuss one of the many static web pages, which represent anti-nuclear groups online.

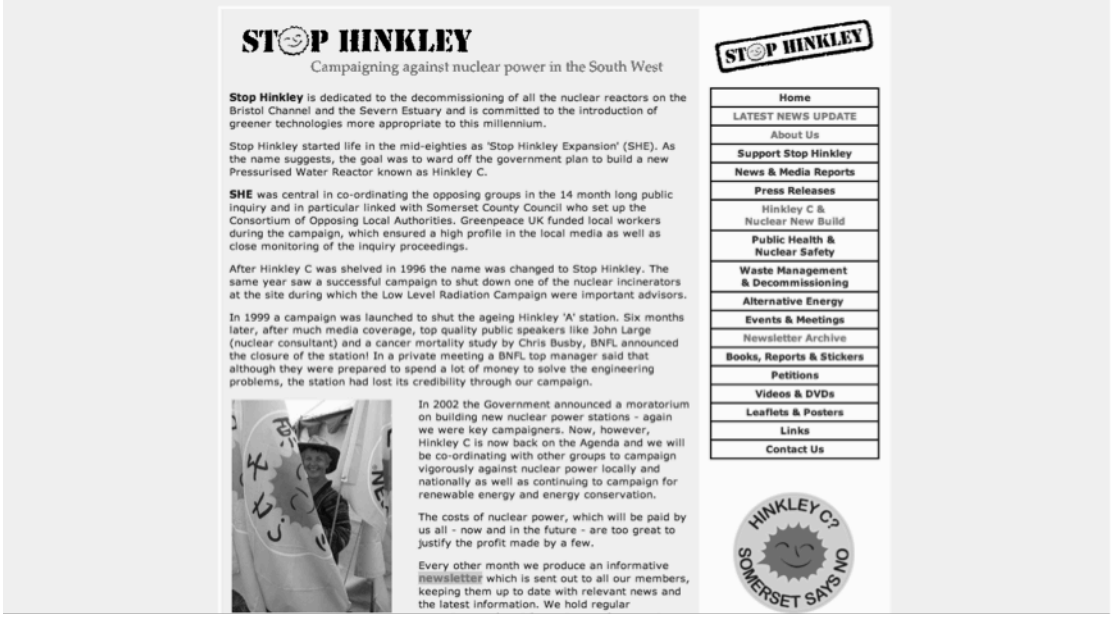


Figure 3. Stop Hinkley: 'About us' page.

In the self-styled 'about us' section of their website, Stop Hinkley, provides exactly the kind of programmatic statements of intent familiar to scholars of collective action frames, which are seemingly absent from certain groups on social media, according the Bennett and Segerberg: 'Stop Hinkley is dedicated to…' Yet even this prepared text reveals that the group has frequently changed identities and objectives in relation to external events and other groups. It was originally called SHE (Stop Hinkely Expansion) when the first attempt to build Hinkley C was proposed, but having

defeated the expansion, it became simply Stop Hinkley as the focused turned to decommissioning the existing B Plant.

Stop Hinkley is entwined in a shifting field of related groups, nuclear power plants, governments and policies and each of these changing associations affects both group stability and identity. Yet this is a very partial representation of the group provided by a 'spokesperson': the static webpage, which gives a rather publicity-facing account of their history. Just as it is crucial to understand how different actors and media articulate a controversy, as a particular configuration of actors, institutions, events, it is similarly important how groups are presented, how they format and articulate the controversy in particular ways.

If we read these accounts in an ANT-inspired way, as actor-worlds, then it is not our concern to explain outcomes in terms of structures or agency, micro-practices or identities but simply to describe the acts of association between groups that are most consequential for the trajectory of the controversy. But so far we have only looked at static very teleological accounts of groups, and must consider the socio-technical arrangements through which texts like these are produced. What happens if we view group formation 'in action' and how can this be accomplished methodologically?

Quantitative and Qualitative Methods

Even if an ANT framework helps to overcome problems of structure-agency or macro-micro conceptually, I argue these problems are also methodological ones: exacerbated by the formatting of the platforms themselves. Just as Wikipedia's data structures both enabled and constrained the researcher's methodological choices, the static web page also encourages certain methodological practices over others – that is the qualitative analyses of the content of statements. It is no surprise that I should discover collective action frames through reading group statements or pamphlets but if I were attending offline meetings, which would be best approached through participant observation, I would be more likely to find micro-organisational practices and resources to be more consequential for group maintenance.

Now on social media, the above techniques are less applicable; because of their scale and complexity, social media platforms encourage the quantitative analysis of traceable interactions at the aggregate level. Yet as I will argue, it is these metrics of

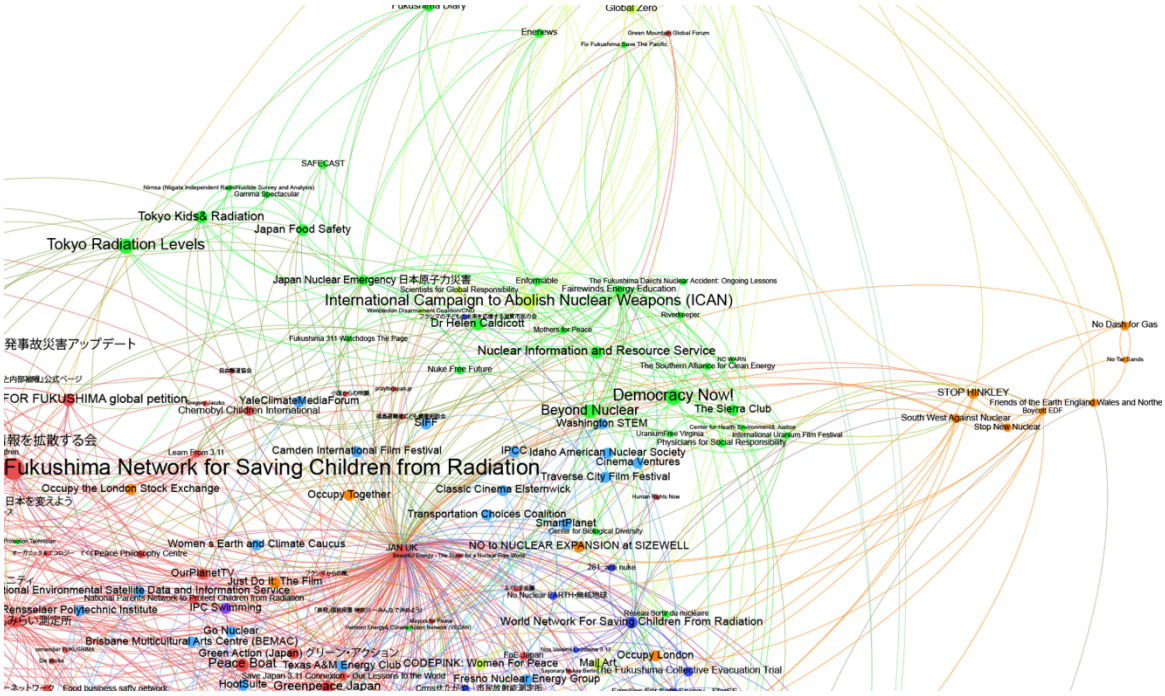
popularity, which need to be interrogated. Questions about group organisation and collective action straddle macro and micro scales and will continue to elude researchers so long as they are kept separate. Similarly reading individual tweets or posts and mapping aggregate trends keeps these different practices separate as well.[131] So how can the researcher deal with these methodological, and thus ontological schisms?

As I suggested in Chapter II, one way to deal with this is through networks, which allow the researcher to zoom from individual to aggregate levels, something which is somewhat facilitated by networked social media (Latour et al., 2012). Facebook provides networked profiles of the sort detailed by Latour et al in which individuals are connected as 'friends' (this is a mutual connection which must be accepted by both parties) but also can be linked to institutions through 'groups' and 'pages'. 'Groups' are private collections of users managed by an admin while 'pages' stand for celebrities, institutions, cultural products etc. which users can connect to by 'liking' them. Specifically in terms of Facebook and activism, Langlois and Elmer (2009) studied a collection of group-pages created for the Canadian election and visualised their shared members in a bi-partite network. They relate this visualisation to the classic issue-network (Marres and Rogers, 2000, 2005), discussed earlier. Networks are one way of addressing the relationality of groups on social media, however, Langlois and Elmer rightly express some scepticism as to whether or not the links in their diagram (created by membership in a group) have the same status as hyperlinks once did. Joining a Facebook group requires much less effort than building and maintaining a static web page for example.

It is not possible to reproduce Langlois and Elmer's methodology for these particular anti-nuclear groups because there are only a couple of infrequently used group-pages representing the nuclear power issue in the UK. Also producing a network of users and pages they like (to mirror Langlois and Elmer's approach) is currently difficult because the Facebook API only freely gives quantities of likes not the names of likers. However something similar can be attempted with *pages* and liking, as the reader will recall from first page of this thesis.

[131] Though it is certainly more complex than this, part of the disagreement between scholars like Bennett and Segerberg and Gerbaudo might be down to the multiple scales performed by different methods: the (macro) quantitative analysis of Tweeting practices versus the (micro) participant observation.

To reiterate, I used the Netvizz application (Rieder, 2013) which takes a given page, identifies all of the page that that page 'likes' and then identifies all the pages those pages in turn 'like'. To make this roughly mirror Issue Crawler, I gathered multiple starting pages and then combined the resulting graphs to see what is common within the network: reducing away the dead ends of the crawl (pages with only 1 like). I started with pages drawn from my knowledge of offline organisations and manual gathering through Facebook searches for 'nuclear power'.



Detail of Figure 1: Like Network https://goo.gl/kZ8AWj: reproduced from Chapter I.

As I described in the very beginning of this thesis, this network increases the list of known organisations and suggests some key relationships. The network is made up of reasonably distinct clusters but, interestingly, these clusters seem *over determined* by the starting points themselves – that is each cluster mostly consists of the interlinked pages liked by each starting point rather than hybrids between the starting points.[132] The exception is the orange group, the rightmost section of which seems to represent the specific UK contingent of anti nuclear groups which bridges the starting points Stop Hinkley and Boycott EDF. These are the groups which are most relevant to the controversy over the particular plant proposal in question which I will address shortly.[133] Liking reveals a very issue-specific network but it also shows how variable

[132] This may be because of the limitation of a crawl depth of 2: a deeper crawl may reveal more connections hiding in the background.

[133] Stop New Nuclear is a container group for many of the local groups on this map but following a Tardian conception, it is on the same level as them (Latour et al., 2012).

and contingent the practice of liking is.[134] As discussed earlier, liking other pages is a completely optional and largely symbolic act, as the individual user will also be able to receive notifications from liked pages independently. Some groups like Jan UK are enthusiastic 'likers' who 'like' hundreds of pages while others, especially the local UK groups, only 'like' a handful. Conventions of liking are simply not as codified as hyperlinking once was in issue networks.[135][136]

The reason I bring this up is that while a like network is a helpful starting point, it is not so simple to import the methodology of Issue Crawler to these new platforms, partly because 'links' are not the same as 'likes' (Gerlitz and Helmond, 2013) but also because I think studying groups 'in action', rather than relational dynamics over time requires a different tack. While the like network presents *external* relationships between groups it presents groups as cohesive entities rather than showing *internal* relationships or tensions between members. Thus it tells one very little about the internal dynamics or shifts in groups or possible organisational practices. Also, there is no rich source of textual content to speak of, other than the names of the groups to study (Langlois et al., 2009). To better understand the contribution of activists on social media it helps to look at the problem through the lens of particular controversies.

In the next section I will therefore examine how various groups through various platforms intervened in the controversy over Hinkley C keeping open this definition of what is at stake in their participation.

02. HOW ACTIVISTS USE ONLINE MEDIA

In March 2013 on the two year anniversary of Fukushima, I attended a protest in Parliament Square, London which brought together groups from across the UK and

---

[134] Interestingly, there is an extreme discrepancy between the number of pages that a page likes and how many users like the page. There is an almost inverse law: pages with a lot of likes, the more corporate or professional pages do not like many pages while the smallest group in terms of in-links Jan UK is the most aggressive liker.

[135] As discussed earlier liking on Facebook has two purposes 1) it allows users to follow the updates on pages they like, which show up in their news feed – allowing them to monitor the page for new content and 2) the page shows up as one of their interests on their profile page, allowing other users to discover pages relevant to them.

[136] Gerlitz and Helmond (2013) have previously compared the politics of liking to previous internet metrics such as hits and links (Rogers, 2004) which were used to determine the importance or authority of webpages.

prominently featured JANUK (Japanese Against Nuclear UK) and Stop Hinkley. This annual event as well as other protests accounted for many surges of activity on various platforms I had been monitoring. Through chatting to attendees and signing up for mailing lists, I gathered a tentative list of groups, web presences, flyers, important actors and events. These starting points (anti-nuclear groups and platforms) allowed me to then look back towards previous controversies and events and track future ones. Looking at specific interventions also allowed me to raise the question of which participatory media were most consequential.

I perhaps naively expected to find anti-nuclear protestors with a huge presence on social media, more or less consistently active, but instead found that the primary online presence of many UK based anti-nuclear groups was on ad hoc blogs, some made through Wordpress templates, static web pages and most importantly email lists. Although Facebook holds more interest for its novelty and increasing popularity, I will briefly discuss a key email list because it was central to organising many of the protests in question, particularly for an older perhaps less tech savvy generation of activists. There is a danger in any study focused on new participatory media technologies that one becomes blind to less trendy and less traceable online settings. Following controversies means taking seriously whichever technologies they relate to. Also email lists provide a good comparison with Facebook pages and prompt similar methodological considerations. However, while I was able to scrape and analyse this platform, I will only be speaking about the list in aggregate form and not in very much detail, not because of technical constraints but rather ethical ones. The list is private – it requires authorization by an admin to view – and while I have spoken to participants, I am not practically able to negotiate access with each of the many users of the list.[137] This is not to say that data visualisations do not have ethical implications nor that the private distinction neatly captures many of the more public facing activities of this list – like many aspects of this project, I think ethical protocols need to emerge from a sensitivity to particular technologies and contingent cases (Rieder, 2013).

Kick Nuclear RiseUp Lists

The London-based 'Kick Nuclear' group acts as a portal for many Londoners who are not directly engaged in a local controversy. Kick Nuclear share some members and

[137] While it would be easy to post a message identifying oneself as a researcher, and asking the group if they were comfortable being studied, there is no guarantee that all members will see this message, or that new members will join after the fact.

promote protests with regional groups like Stop Hinkley and Radiation Free Lakeland, which are all part of the larger alliance Stop New Nuclear. Kick Nuclear's main web presence is in the form of a blog and email list on activist website RiseUp Lists (riseup.net). RiseUp is a subscription service: one's request to join must be approved by the admin, while Facebook pages are open.

Image 4. RiseUp Email Lists interface. Messages are also forwarded to users' email. Message content and user names have been redacted from the image above.

RiseUp Lists are composed of posts by a users and replies to those posts, visualised as a thread when the user copies the message title and adds 'RE:' This format of the archive is the particular way in which the mass of data is made sensible and legible to users. However there is nothing like Wikipedia's tools or bots, let alone ready-made Digital Methods for these email lists. Again there is a variable traceability, which might incline certain researchers to study email lists in certain ways or more likely ignore email lists all together because they are hard to scrape and visualise. So how can this format be studied in a way which cuts across the micro-macro, quantitative / qualitative divide?

One potential solution comes from a paper, which is very old in internet time, Warren Sack's 'Discourse Diagrams' (Sack, 2000). Sack proposes an approach for studying the similarly structured internet News Groups, which he refers to as 'Very Large Scale Conversations', which almost seem quaint by comparison with conversations on online platforms. The tool consists of three panels: on the left is a social network graph, which shows which users interact with other users by commenting on the same post: if User

A responds to User B then A and B are connected. On the right pane is a 'semantic network' of the topics discussed. In this process all nouns in messages are profiled for which verbs, adjectives and other nouns they regularly appear with. Nouns with similar profiles are seen to be related in that they can be used relatively interchangeably in statements. These relationships are visualised as a hierarchical network diagram for each word with more related words branching out into less related words. Finally, at the bottom there is a diagram of the structure of each threaded conversation – how many replies, how many replies to replies.

Image 5. Warren Sack's Discourse Diagrams

Together the first two diagrams relate individual practices, acts of participation, such as commenting, with the *content* of that participation, which is an advantage given the methodological split identified earlier. But there is a danger with keeping these diagrams separate. The social network diagram on the left seems to materialise social relations as if they are divorced from the medium which enacts them. If read uncritically one might forget these are merely traces of comments and replies. The semantic network on the right presents the danger of abstracting *content* from the infrastructure through which it travels, individual posts, as if one had direct access to the group or a sub group's collective thoughts. Both problems are somewhat addressed by the existence of the bottom graph which shows the material structure of threads, but this means oscillating between different abstractions rather than highlighting relationships between them.

One way to keep the actual data traces in the picture would be to combine the left and right networks with the bottom as two bi-partite graphs, a type of visualisation I discussed in Chapter III. First I collected all posts from the year 2011-12 of the Kick Nuclear list by placing the list in 'chronological view' and scraping it using the Google Chrome extension Scraper (which I used to parse data in lists using the HTML tags in the page). Starting with the graph on the right, I made a simplification of the semantic network. Semantic networks are based on formal properties of language, which may have a bias toward certain official modes of speaking and of course certain languages, but the main problem is that such a method is not tailored to the specifics of the medium – how would, for example abbreviations like RE: and FW: be understood semantically? Instead I used a bi-partite co-occurrence network produced using ANTA (Venturini and Guido, 2012). Essentially when words or phrases appear in the text of a comment together, they are connected.[138] The more often words co-occur in posts the more related they are presumed to be as entities.[139]

[138] Although ANTA is much simpler than a semantic network, it is not without its own assumptions about what is most important. ANTA uses stop lists to remove common words like 'a, and, the' and also queries a database called Alchemy, which identifies and resolves different spellings of known proper names. It would be preferable in future to use a type of co-word analysis which does not assume the importance of words and merely uses a 'stop list' to remove common words like 'and', 'the' etc. The current visualization is limited because ANTA does not recognize enough of the names in this topic area.

[139] In future, it would also be preferable to use the full text of the messages but this has been difficult to scrape, so instead I have used only the post title. The problem is that titles in threads then have the same basic text and will naturally cluster together – it would be better to group threads and then visualise the number of replies as node size. The strings '[Kick Nuclear]' and 'RE:' have been removed because they appear in too many messages, but other artefacts like 'fwd:' have been left in because they suggests which content comes from outside groups. Below I have produced this graph using posts from the years 2012-13.

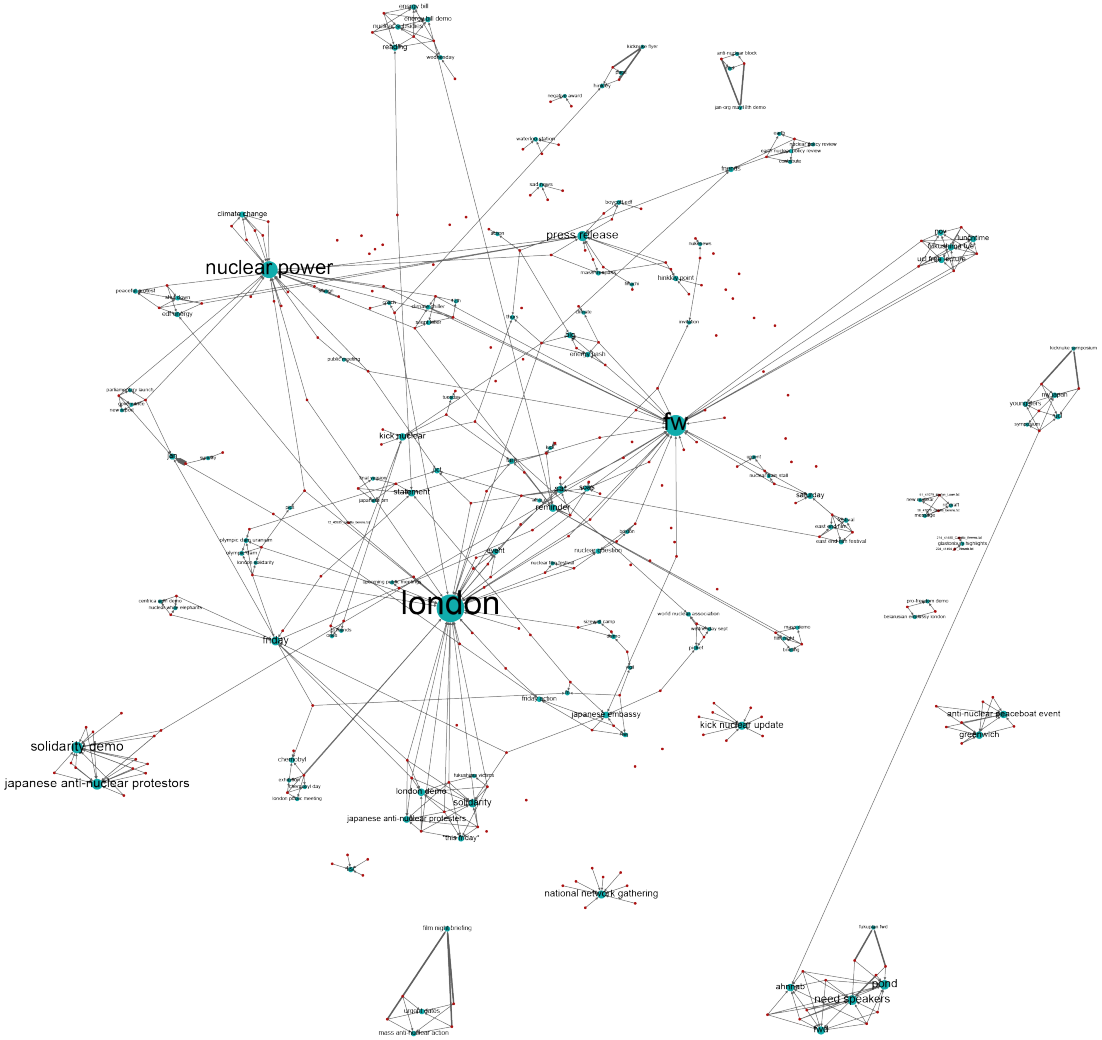


Figure 7. RiseUp Word-Post Network https://goo.gl/N36JSA: Bi-partite Co-word map created with ANTA. Grey nodes are threads or single posts and blue nodes are words, both sized by degree – number connections. Spans the years 2012-3.

This network gives a quick snapshot of how the list is used in relation to nuclear controversies. The email list appears to be very focused on on-the-ground events, judging from key words like demos, solidarity, gathering, action, dates, protestors, film nights, speakers, lecture and symposium. This also explains the prominence of the node London as the main location where these interventions occur. The clustering of the diagram, using the algorithm Force Atlas 2, however, shows relatively little about themes only that they are focused on particular discreet events or campaigns. Although there is not space to dwell on this point, this map gives a sense, in a similar way to the like network, of some of the objects, controversies, settings and other groups through which the group / list obtains its shifting identity, this time visualised through textual outputs rather than liking. However, I cannot claim much about these associations without seeing how they are performed through posts themselves. Are these proposals

and connections accepted by the group or other groups passively ignored? It is important to look at the second bi-partite map to get a view of some of the internal dynamics of the site. The second network, follows the same logic as the network of interactions (the left of Sack's graph) but includes the posts themselves – users are connected to posts they comment on and the more times they comment, in the case of a long back and forth conversation, the thicker the lines become.[140]

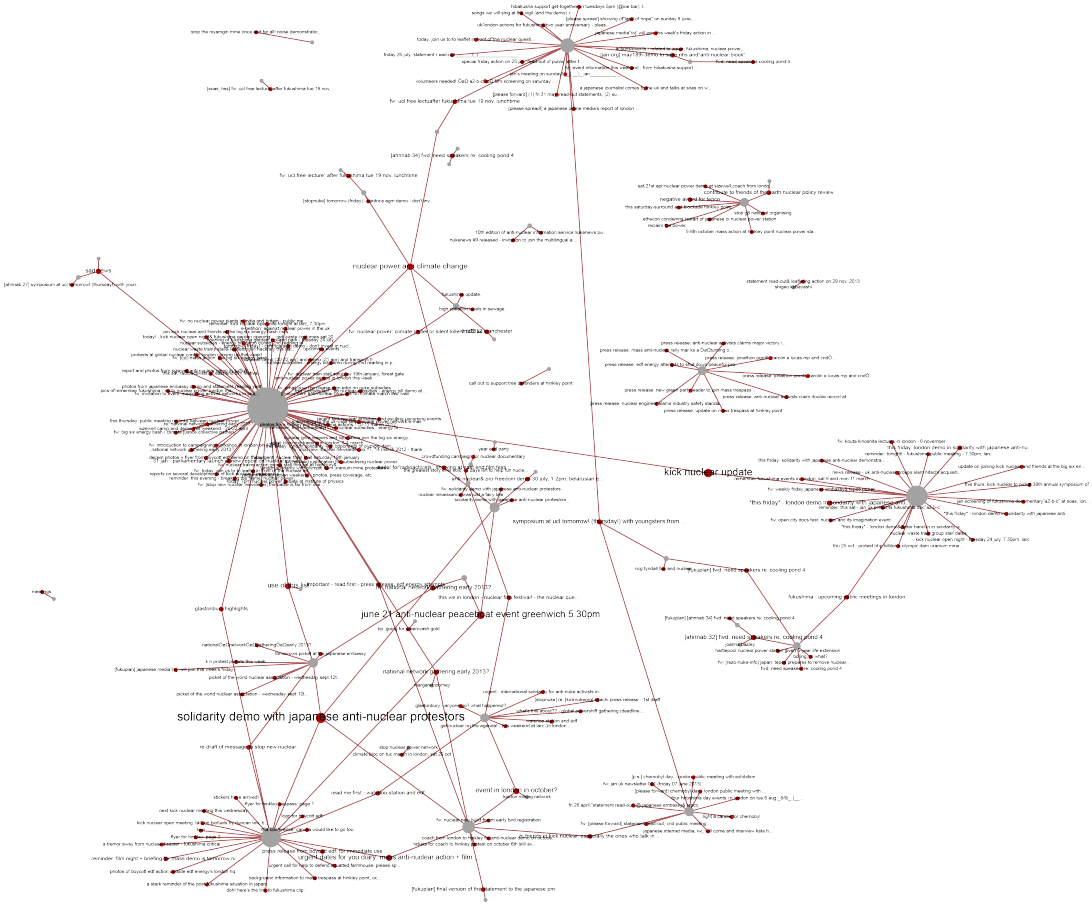


Figure 8. RiseUp User-Post Network https://goo.gl/Zg0ILZ: Bipartite network created with Table2Net. Red nodes are threads or single posts and grey nodes are users, sized by number of posts in thread or number of posts replied to by user. User Names have been redacted. Spans the years 2011-2.

Looking at this User-Post network one can see immediately that most posts are not replied to as signified by the halos of single posts centring around certain users.[141] This

[140] This unfortunately removes the structure of threaded posts (the bottom graph in Sack's diagram) but this can be simulated with node size to indicate number of replies. The maximum replies to a conversation is around 10 and many posts are not responded to – they are only connected to one user, the original poster.

[141] This of course does not mean that they are not read by people on the list or have important effects in terms of organising on the ground, merely that they have not elicited a reply.

suggests that the list is mostly used as a conduit through which information passes – not a place to linger and comment. This is a relatively distributed network of users: although there are a few key posters who are responsible for most of the content – including one who is the group admin. Interestingly the main posters rarely engage with each other's content, which makes sense because they seem to represent leaders or press officers of different organisations who use the list.

Just as in the last chapter, it is helpful to start with frequency of activity as a rough measure of controversiality, which then must be interrogated through qualitative analysis to determine if these pertain to controversies or to medium-specific attributes. However in this particular case the most replied to posts mostly consist of logistical discussions about the planning and organisation of events. Another type of post, which does not receive many replies, but is far more interesting, are information sharing posts, mainly presentations of links to relevant news stories. These posts give key information to group members and also may be occasions to reassert the key goals and interests of the members.

However, the main point I wanted to make using this list was a methodological one: that we need both networks together to understand the contribution of platforms. The co-word network, in some sense stands for some of the external relationships between the groups and other groups, but one cannot grasp the strength of these connections without getting a sense of which of these external relations are controversial within the group or which are taken up. Reading across both, but also zooming in on the individual posts in question showed that email lists are predominantly used for organising offline actions but also for sharing content and news stories, both of which have implications for the group and the formatting of the controversy by that group.

In this way, email lists reproduce much of what we know about activism and issue-networks (the members of which also made use of email lists). They are used to organize offline protests and coordinate interventions (Mercea, 2013) but they are also *informational*, circulating specific articulations of issues (Marres, 2006). But will the same be the case with the more publicity facing technology of Facebook.

Facebook and Bridgewater

On Facebook, I will start with the page for the group Stop New Nuclear who, as mentioned earlier, are an umbrella organisation for most of the relevant groups involved in the controversy and was identified as well-connected in the 'like network'. In the two years following Fukushima, their coordination efforts were pivotal in organising three direct actions at Hinkley point, the proposed site of a new power plant, with an association of other groups. As with the Kick Nuclear List, the Stop New Nuclear Facebook page is an open space for other groups to cross post and coordinate which makes it a good vantage from which to observe inter-group dynamics. The Stop New Nuclear page below shows a selection of posts in reverse chronological order. How these posts appear in the 'timeline' of users though has to do with Facebook's proprietary algorithm which selects relevant posts, for example, based on high engagement (number of comments and likes and shares). The page admin can also supress posts by other users in 'Posts by Users' section in the left. By scraping the page with Netvizz (Rieder, 2013), it is possible to obtain all the posts, not just the ones currently on display.

Image 6. Stop New Nuclear Facebook Page.

Facebook pages are similar in structure to email lists, they are composed of posts by users and (semi) threaded responses and replies, but there are several important differences: one of which is 'likes'. Likes, as discussed earlier, are a show of enthusiasm

for a post which, depending on user settings (potentially) places that post into the newsfeed of friends of the liker. 'Sharing' is the act of re-posting content directly on ones wall, again potentially entering the news feed of one's friends.

While email lists require subscription and vetting, 'liking' them on Facebook is not synonymous with being a member of the group or the endorsement of the objectives stated on group websites. The Facebook page may be 'liked' by members of related groups, disinterested citizens, journalists, corporate spies or researchers.[142] While the RiseUp lists are to some extent guarded and internal, Facebook pages and Twitter of course are public, inviting comments from anyone, even group opponents. Facebook pages are open in the sense of Wikipedia being open but the extent of openness or visibility is again a joint accomplishment of platform and user activities, as will become clear later.

These differences in formatting in relation to email lists can be clearly seen by producing another 'user-post' network. For the following Graph, which was again made using the Netvizz scraper and Gephi, the lines between users and posts correspond to commenting on, *or* liking the post.

[142] I in fact discovered one more researcher other than myself, in the following post:

> We_re posting this questionairre up on behalf of a student from Plymouth University who is doing a dissertation on Hinkley Point and nuclear energy…

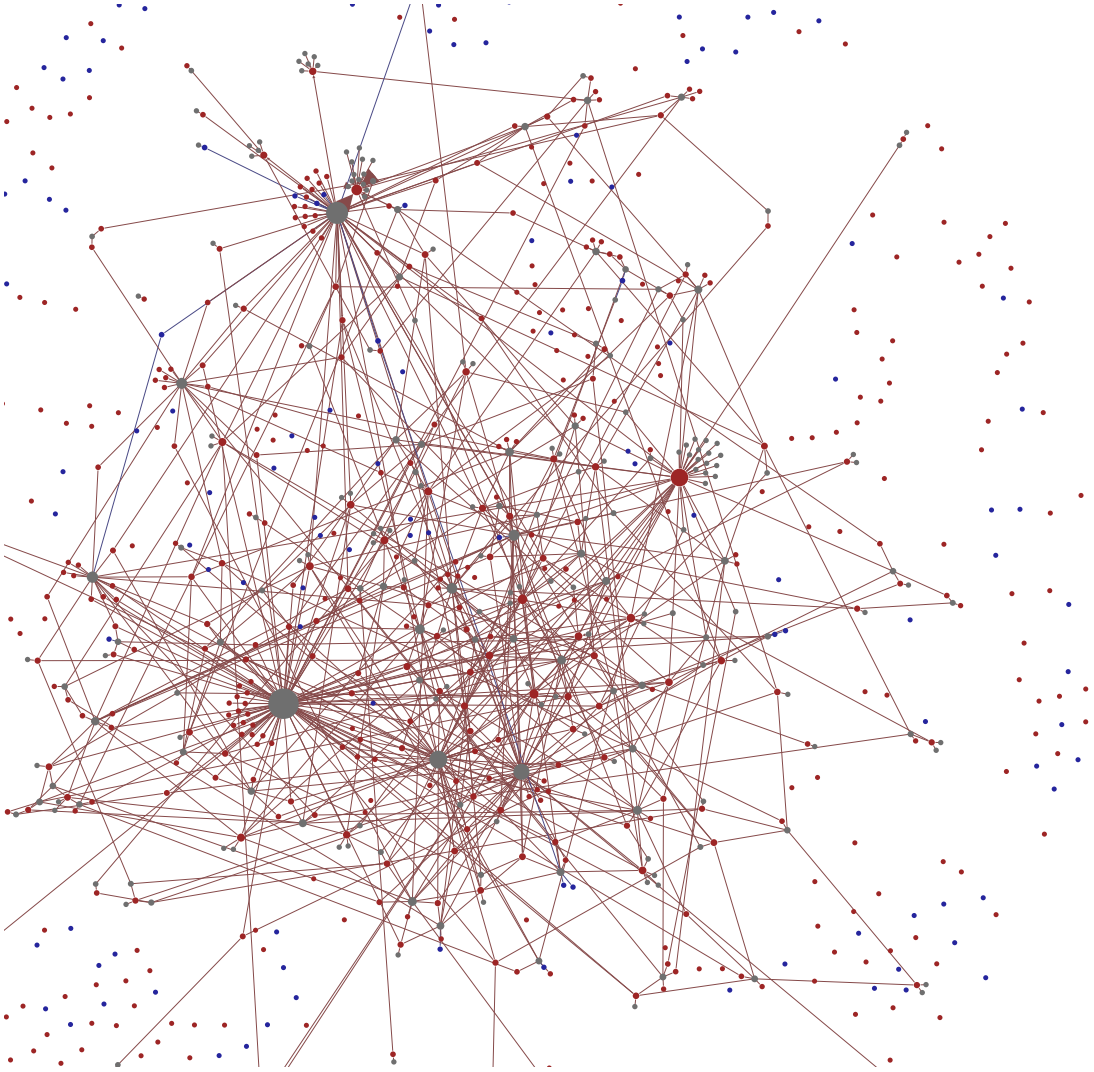


Figure 9. Stop New Nuclear User-Post Network https://goo.gl/li7EQP: Bipartite network created with Table2Net. Red nodes are threads or single posts and grey nodes are users, sized by number of posts in thread or number of posts replied to by user. User Names have been redacted. Spans the years 2011-2.

Much like the email list, the majority of posts originate from about four users including the admin account (e.g. user: Stop New Nuclear), which is a special account operated by certain users with access to the page. Compared to the email list, however, this map is more of an undifferentiated mess. This is partly because there are more users and more total connections, which one might attribute to the ease of participation – the ease of liking or making short comments as opposed to the (comparatively strenuous) labour of posting an email list reply. For this reason it would be better to differentiate between likes and shares and comments, because they entail different activities, but this is not currently possible given the Facebook API. In any case, it is hard to tell

anything about the clustering of this content precisely because of how dense the graph is. I will return to this problem later in the chapter.

These traces of activity are easily quantifiable and graphable and available from the Netvizz interface at the push of a button. But what about the less easily scrapable *content* of the posts? Using a similar procedure to the email lists, I first scraped the posts from Stop New Nuclear from 2011 to the end of 2012 using Netvizz and then brought the post text as separate files into ANTA.

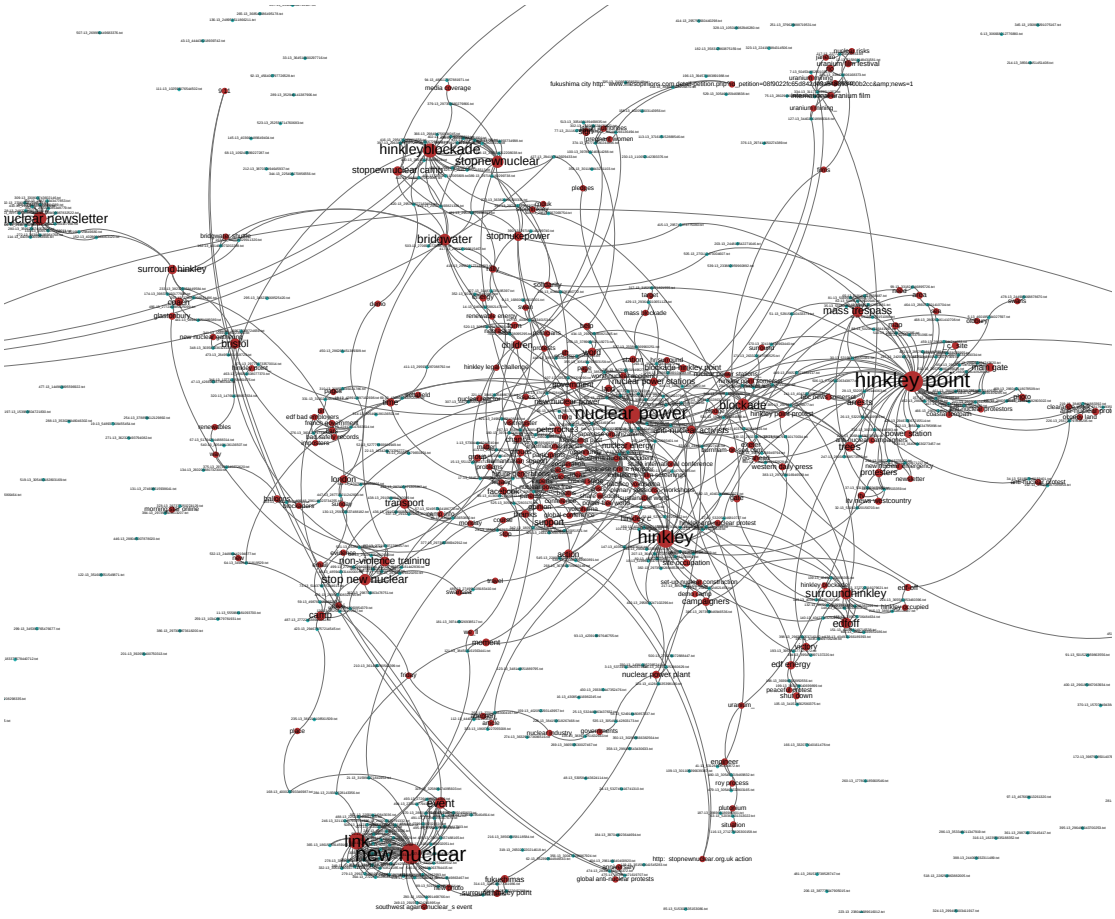


Figure 10. Stop New Nuclear Word-Post Network https://goo.gl/uALalV: Bi-partite Co-word map created with ANTA. Grey nodes are threads or single posts and blue nodes are words, both sized by degree – number connections. Spans the years 2011-12

As the map shows, this Facebook page is similarly focused on action on the ground but most of it refers to a very specific cluster of actions variously known as the 'Hinkley Point Blockade' or the 'Hinkely Occupation'. This event in October 2011 was conceived of before the Fukushima disaster but gained extra importance in relation to it. One very prominent node is thus the location 'Bridgewater' where the camp will be held with

references to shuttles moving people there. There are sub clusters related to 'training' for the event in relation to direct action and law seminars.[143]

These posts mainly cluster around particular interventions. There are also 'periods of latency' (Melucci, 1996) between events and it is interesting that in these gaps the admins keep the momentum going by, as with the email lists, sharing news stories, blog posts and videos.

> Some interesting info rising radiation readings on US west coast as a result of Fukushima. http://www.youtube.com/watch?v=SmsFfXXkPU4

These are not about protests but about knowledge claims, which are intended to inform group members and also be spread through their networks by being 'shared'. But there is a third type of post, which is neither strictly organizational nor informational, the best example of which is the largest (most liked or commented on) post, according the user-post graph.

> Fancy annoying nuclear power company EDF today? & everyday till Halloween? Of course you do!! Their flagship sponsorship deal The London Eye is asking what your biggest fear is... er_ EDF building new nuclear power in the UK... Fukushima_ childhood leukemia_ you know the kind of thing... […].

It proposes direct action not on the ground, but in the purely online space of Twitter. The effect is similar to asking people to show up and hi-jack a public hearing or write letters to their MP but this time it targets a social media space, ruining a carefully orchestrated PR stunt. One might call these interventions a 'virtual sit-in', in the sense that it is marshaling followers to occupy an online space but this term has come to refer more specifically to Hacktivism, to denial of service attacks or interventions which are intended to obstruct or sabotage.

[143] There are also a few smaller clusters which are artefacts of medium specific types of posts. 'New Nuclear Update' is the phrase which appears in several posts directing users to a newsletter and 'New'- 'Nuclear' - 'Link' pertains to automated posts 'No New Nuclear shared a link' which is automatically generated when URLs are posted. There is also a cluster which refers to a series of news stories, shared on the page, about EDF's legal scandals – so there is also some select link sharing going on.

03. DIGITAL DEMONSTRATIONS

A more appropriate term might be 'digital demonstrations'. This calls to mind Andrew Barry's (2001) twin use of the term 'demonstrations', which refers to both protests (civil disobedience, sit-ins) *and* the scientific demonstration or proof in public, such as anatomical lectures in the age of amateur science. Barry draws attention to the fact that most political protests have an 'empirical' dimension to them: '…to conduct a political demonstration can be a matter of making visible a phenomenon to be witnessed by others.'(2000: 178) This could be highlighting an injustice or corruption or presenting alternative scientific claims (as is often the case in nuclear disputes Nelkin, 1974; Nowotny and Hirsch, 1980) but also merely demonstrating that a significant opposition or public exists (Marres, 2012a).

In the particular case he examines, a protest over the construction of a road through a nature reserve, the assembled protesters attempted to reveal environmental destruction 'as an emerging reality'. This involved carnivalesque protests, interventions by artists like Christo and filmmaker Werner Herzog and the occupation of trees marked for destruction, all with a view to courting both alternative and mainstream media. In STS studies, when science or expert institutions are involved it is perhaps obvious to view activists as producing alternative knowledge: 'experiential knowledge' (Rabeharisoa, 2003); 'research in the wild' (Callon et al., 2001); 'evidence based activism' (Rabeharisoa et al., 2014) etc. But Barry's argument gestures to activism more generally. The point is that while politics has proliferated in a variety of settings[144] and in a variety of forms, it is not everywhere; politics and publics must be *made visible* and this is an accomplishment of a variety of actors, especially media.[145] I will call *digital* demonstrations, actions which are predominantly online intervention aimed at making things visible: both in terms of (discursive) articulations and through modes of assembling, including the accumulation of likes. Digital demonstrations could be a way of understanding of 'clicktivism' in positive terms rather than an impoverished form of offline activities. It also allows a more generous reading of what

[144] Barry prefers to separate 'politics', defined as 'spaces of disagreement' from 'the political' as in the institutions of politics which can just as easily be responsible for the silencing of spaces of disagreement or what he calls 'anti-politics' (Barry 2005).

[145] In the case of the road protest, the mutually affected 'public' was a very heterogeneous group which spanned the political spectrum and even disagreed on basic strategy. Ironically, Barry claims that this was to the group's advantage because they could not so easily be painted as NIMBYs (Not in My Backyard) or ideologically environmental and written off by the media.

counts in terms of public issues above and beyond more narrowly understood science controversies.

One point in invoking this empirical dimension of protest is that demonstrations, whether making knowledge claims or materializing an opposition, can and do fail, their claims are dismissed or the plight of participants ignored. They have, if not implicit truth conditions, then implicit criteria for success.

Contesting Interventions

Returning to the Facebook Page and now zooming in on the text of the posts: the attack on the EDF Twitter handle is met with resistance. First, one pro-nuclear user criticizes the child leukemia claim, demanding the page admin produce some 'facts', which attacks the content of the original post. The charge of irrationality is a typical tactic used to silence anti-nuclear activists (Nelkin, 1974), this is despite the fact that both sides later engage in a rather high level debate about the statistical methods used in the UK COMARE report into childhood leukaemia versus those of the German KiKK study. This is a contest of competing knowledge about nuclear, but it is also crucially about what kind of knowledge is admissible in debates such as these:

> User B: How about a scientific fact that backs up anything you've said about nuclear energy (that is what this page is about, right?)?
>
> Admin: Fact: being a nuclear physicist doesn't qualify you in biology. Fact: the oddly named 'health physics' is a way of thinking that attempts to impose abstract mathematical models onto the way that radiation interacts with living things such as tissues and organs etc.
>
> User B: Good luck with your campaigning. Please know that there are plenty of knowledgeable people who will continue to interfere with facts to dispute your emotional pleas.
>
> Admin: …emotional pleas are also important. We have other options. Our consumer life styles are not sustainable. Emotional and social intelligence are so easily pushed aside as we destroy our planet for future generations.

Because posts are circulated to a potentially wide variety of friends-of-friends,

Facebook enables quite open confrontations between opponents in the nuclear issue in which the boundaries of what counts as science / non-science or rational / emotional are actually discussed. After a few such attacks from the pro-nuclear side, the page admin responds, 'There seems to be some campaign going on here….'. The implication is that these pro-nuclear voices are not innocent members of the public but an example of 'astroturfing' — or the engineering of seemingly grassroots activism by a private company who pays people to set up fake accounts posing as members of the public (sometimes known as a click farm).

> Admin: Definition of 'troll': s someone who posts inflammatory, extraneous, or off-topic messages in an online community, such as an online discussion forum, chat room, or blog, with the primary intent of provoking readers into an emotional response or of otherwise disrupting normal on-topic discussion. And that's a fact!

Interestingly the Stop New Nuclear admins see this invasion of pro-nuclear voices on their page as 'trolling' but does not apply the same logic to what they are attempting to do to the @londoneye Twitter account. One of the pro-nuclear users draws attention to this:

> And isn't this post about annoying (or 'trolling') EDF to begin with? It seems to be an encouraged practice from what I can tell.

The pro nuclear users are of course also engaging in a digital (counter) demonstration but there is a fine line between digital demonstrations, astroturfing and trolling because they all involve potentially occupying an online space which is not their own and making their presence felt. Astroturfing, which it should be said is very hard to definitively prove, involves an element of deception and far greater financial resources. Trolling could be seen as causing trouble for sport rather than for political gain, though the case of 'hacktivist' collective Anonymous certainly blurs this (Coleman, 2011). The point is that what counts as 'proper' or valid participation and political action, what register political participation is understood in, is a topic which is debated by the participants within Facebook, but also as I will show later, partially defined by the platform itself.

On 5 October, there was a flurry of links recapping the media coverage of the Hinkley Blockade:

> We also made it into Euronews in German_ French_ Italian_ and English! Here the German version: http://de.euronews.net/2011/10/03/atomkraftgegner-demonstrieren-in-hinkley-point/

This coverage curated by the Stop New Nuclear, is mostly in smaller local papers or nuclear focused sources. The Morning Star article claims there were 200 protesters (The BBC reports 'up to 400') and makes a special point of most of them being local. Other sources lead with the line '7 arrests at nuclear protest' (or 6 depending on the source). One of the main ways media intervene in the fate of social movements is by exaggerating or downplaying numbers (Gitlin, 1980).

In several stories, the word 'symbolic' interestingly comes up in relation to the protests, often through quotes by activists. This term on one hand is meant to obviate the need for the protest to 'work', in the sense of the blockade actually halting energy production, but also to draw attention to the more theatrical gestures at the camp. For example at the aforementioned Hinkley blockade, the group released 200 balloons from the site, 'symbolising' the potential release of radioactive fallout from a nuclear accident. The balloons were tagged so that anyone who found the balloons could email their eventual location back to the group. These locations were placed on a map which was circulated on Facebook, Twitter and the RiseUp lists.[146] But this evocative and poetic statement was met with some distain from outside the group as in the following posts which were actually cross posted on Twitter. [147]

[146] Available at: https://twitter.com/StopNewNuclear/status/121660283298119680 (Accessed 5 July 2015)
[147] Note that the user @Pro_fission changed their handle to @YesNuclearUK since this post: https://twitter.com/StopNewNuclear/status/121666756044525568 (accessed 5 July 2015)

Image 7. Stop New Nuclear Hinkley Occupation October 2011

Many of the posts by Stop New Nuclear contain Twitter type artefacts, such as hashtags because the moderator was no doubt using a cross posting application like Dlvr.it. By posting on Twitter, Facebook and their own blog, the audience, and thus the possibility for disagreements, becomes much greater.

Image 8. Tweets from the Stop New Nuclear Account

It is ironic that this action is judged in terms of what it 'proves' in the language of science, as opposed to what it so effectively evokes as a symbolic gesture. It begs the question: what kind of action would it take to convincingly prove something?[148] One could judge this intervention based on its resemblance to a scientific experiment – a knowledge claim, or it could be judge on the number of people it reaches – the number of balloons returned, but I think this symbolic gesture cannot be captured by either. Similarly are we to judge the success of digital demonstrations purely in terms of likes? It is this inclination towards quantitative evaluation, which I will consider next.

The Quantification of Participation

Taina Bucher (2012) claims that Facebook produces a regime of visibility / invisibility.[149] She has in mind here older broadcast media which, through framing,

[148] There was a similar controversy, described by Collins (1985) over an event organised by the nuclear industry where a spent fuel flask, normally transported on trains, was subjected to a 'worse case scenario' train crash, which it survived. Of course Greenpeace an other organisations critiqued the conditions of the experiment, which assumed crashes would happen in certain ways. The action was seen as definitive without it being repeated or other possible crash scenarios being explored. Collins makes an analytic distinction between 'experiments' by a core set of scientists and 'demonstrations' or shows of virtuosity aimed at the public and showed how the nuclear train performance confused the two. Rather than maintain these distinctions which have to do with inside and outside of science, it is I think preferable to recognise that the way these interventions are evaluated is at stake in the particular intervention.

[149] Bucher reverses Foucault's panopticon concept, in which subjects are attempting to hide from an all seeing gaze, in Facebook they are actively courting their own surveillance, in

gatekeeping and agenda setting, set the conditions and the terms on which information became available: '…becoming visible, or being granted visibility is a highly contested game of power in which the media play a crucial role.' (Bucher, 2012: 1165) But as discussed in Chapter II, now these processes are also governed by algorithms as part of socio-technical assemblages (Gillespie, 2014).

Each item on Facebook – a post, an image, etc. are considered 'objects' and interactions with objects (comments, shares or likes) are 'edges' in the lingo of network analysis. Facebook's Edge Rank algorithm determines which objects should appear in others user's news feeds by ranking interactions by 1) affinity with other users 2) weight of the edge (number of comments etc.) and 3) time decay.[150] These enshrine Facebook's assumptions about what is relevant to users and creates a popularity game (which has spawned a whole new industry of 'News Feed Optimization', like Search Engine Optimization), a game which is increasingly quantitative or calculative.

As Gerlitz and Helmond (2013) describe in a related paper about the proliferation of Facebook's like buttons, 'In this Like economy, the social is collapsed with the traceable, as user affects and interactions are instantly measured for data mining purposes and multiplied in order to generate more traffic and engagement' (2013: 4). Although it may be an obvious point, clicking like overwrites a variety of responses including excitement, agreement, tacit approval and even sarcasm.

And yet, 'likes' cannot simply be disregard as 'false', failing to capture the richness of social actions because as Gerlitz notes elsewhere (Gerlitz and Lury, 2014) these rankings *format* behavior, orienting users towards future forecasting and even influence the temporalities of online activity. As they put it, rankings on social media are not a measure of participation, they are a 'participatory measure' (they elicit participation). They are part of the empirical reality being studied not sitting outside of it.

Helen Verran (2012) who studies the performative effects of numbers as semiotic-material objects, distinguishes between uses of numbers as iconic, symbolic and indexical, following Pierce's semiotic typology. Iconic numbers come to name

---

contrast to the Foucaultian analysis of surveillance, in which the constant possibility (but not reality) of being watched creates conformity.

[150] In which case 'liveness' and 'realtime' are privileged (Marres and Weltevrede, 2013).

something and symbolic numbers imbue what they name with value, which is one way of understanding what accumulated likes and rankings do to social life. But she continues:

> 'Engaging with numbers indexically involves explicitly working with what using them as icons blithely takes for granted, and using them as symbols insistently denies: the need to wrestle with the always and already over-whelming, blooming, buzzing real.'(Verran, 2012: 120)

In the next section I will make a modest attempt to re-embed these quantitative traces by placing them in dialogue with more qualitative data.

04. TRI-PARTITE GRAPHS

My suspicion that this quantification of participation has implications for the balance of methods we use. In interrogating the idea of clicktivism it may be reductive to simply run quantitative metrics on the clicks, these digital traces of participation, as if they can so easily disclose the activities and intentions of actors and groups. This means potentially buying into the same logic of 'popularity' or 'liveness' inscribed in these platforms (Marres and Weltevrede, 2013). [151] And yet, as just noted, they also elicit and shape possibilities for action.

As I proposed in Chapter III, these capacities of quantitative and qualitative techniques need to be redistributed: we need to *qualify* these quantitative traces and make the incalculable more *calculable*. To this end I will propose a data visualization which combines the easily quantifiable and the less quantifiable. Earlier in this chapter, I used two bi-partitite networks, one of which represented the easily quantifiable traces of interactions between users and the other which, crudely, visualized the harder-to-quantify, unformatted text of the posts.

However, one of the main problems with networks is that they tend to flatten time or at best appear as a sequence of static time slices, which can cast very fleeting associations between entities as stable connections. Also, time, the focus on new content and

[151] One reflexive response (Rogers 2009, 2013) has been to repurpose these traces (likes, rankings, links, shares) and use them to reverse engineer the platform. Rather than taking this data as a transparent window on to the social, like a futuristic opinion poll, through comparison or studies over time, Google rankings can be shown to have bias or national differences, revealing how they format social life.

'liveness' (Marres and Weltevrede, 2013) is part of the effects of the platforms and also must be interrogated. Finally the reciprocal influence between likes and textual contents can only be seen by looking at changes over time.

My imperfect solution involves combining the two networks, one containing users and posts and one containing words and posts, by joining them at the posts and then arranging the posts vertically along a central y-axis in time order by timestamp. I will refer to this graph as a tri-partite network because it contains three types of nodes. I then let the user-nodes on the left and the word-nodes on the right reflow using another gravity-based spatialization algorithm so that they settle closest to posts in which they appear most often. It should be noted that this method of reflowing is not very satisfactory, the placement on the page is somewhat arbitrary – nodes with regular engagement are inclined to be drawn to the middle, regardless of where their temporal focus is.[152] Perhaps in future the vertical placement of words and users could be based on another logic, which is more meaningful, but for now I will give the nodes a horizontal property 'degree', simply the number of connections, which distinguishes actors who comment on particular posts and those who merely 'like' everything equally. Similarly, this distinguishes between generalised words used frequently and words which are more specific to certain posts. Nodes with a higher degree are placed further away from the graph and nodes with a lesser degree are closer to the few posts they interact with. Nodes from the actor column and posts are resized based on the frequency of engagement (both likes and comments) and words are resized by the number of posts containing them.

[152] Also posts which are long and thus have more words, may also weight the graph in somewhat arbitrary ways.

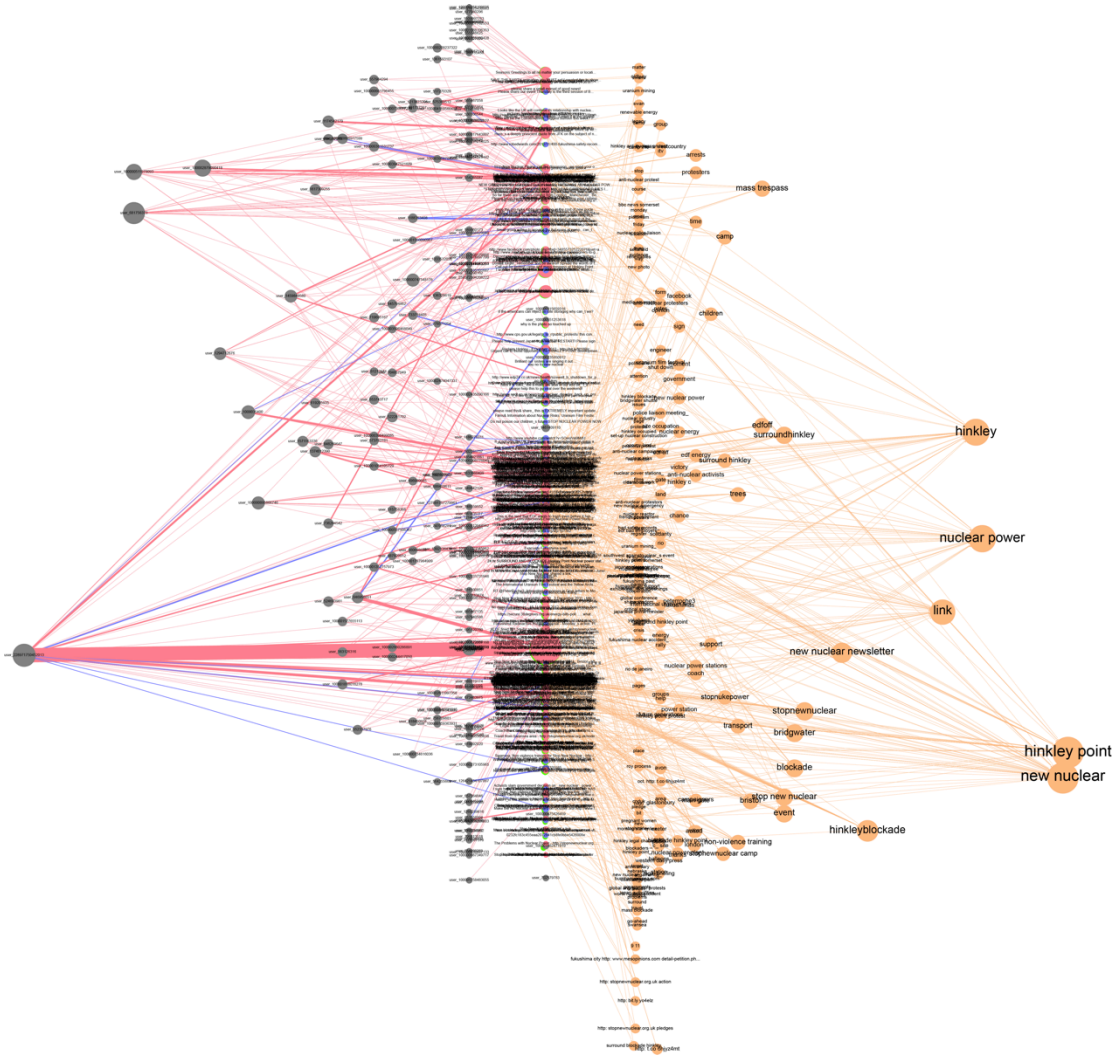


Figure 11. Stop New Nuclear Tri-partite Graph https://goo.gl/qTMIH8: Each graph is composed of three columns (users on the left, posts in the middle, words on the right). The posts are coded red for page or admin-generated posts and blue for user-generated posts, this is also true for the lines connecting users to posts.

This graph can be read by scanning vertically each of the three elements from bottom to top and then interrogating points of interest horizontally across. What it quite simply shows is the relationship between discursive content and who engages with it. Most other social media studies would have, by various other means only looked at the patterns of participation on the left side of the graph which are the more easily quantifiable and analysable while other researchers might look at the content of individual posts without getting a more macro sense of patterns or rhythms of engagement. The other thing this graph allows the researcher to do is to more easily escape the top end of frequency measures or 'liveness' of likes: for example, highly connected actors or frequently used terms can now draw one's attention to posts with perhaps low engagement or periods of relative inactivity.

Reading the middle axis of the graph, the posts appear to come in spurts (when the posts bunch up and the text starts to overlap and become unreadable). Linking these spurts to the terms on the right gives a quick indication of what prompts them. Starting from the bottom, just after Fukushima in April 2011, there are a few isolated gatherings, which correspond to the initial announcements of the Hinkley Blockade in October. These are leading the event attempting to build momentum. The page really explodes as the event approaches. There are 41 posts directly related to the camp as it happens between 30 September and 3 October. This dense bar of posts a third of the way up the graph contains everything from messages of support to requests for carpooling.[153]

But interestingly, the largest post in terms of quantitative engagement (looking at the concentration of edges on the left side of the graph) is the 'digital demonstration' on Twitter and the debate that ensued. It should also be said that the 'engagement' expressed quantitatively masks the ambivalence and the outright hostility of the comments and responses. But if one possible outcome of this is spreading messages through the network then such a response, which might be favoured by the algorithm, has a greater chance of spreading.

[153] The formatting of the posts reveal that many of them are simultaneously sent from a static blog and the presence of #'s reveals that they are also sent out on Twitter. So Facebook is not necessarily the primary means of disseminating this information, it may specifically target a different demographic of less proactive but interested followers (see Mercea 2013 discussed earlier) who may or may not attend meetings or even the protest.

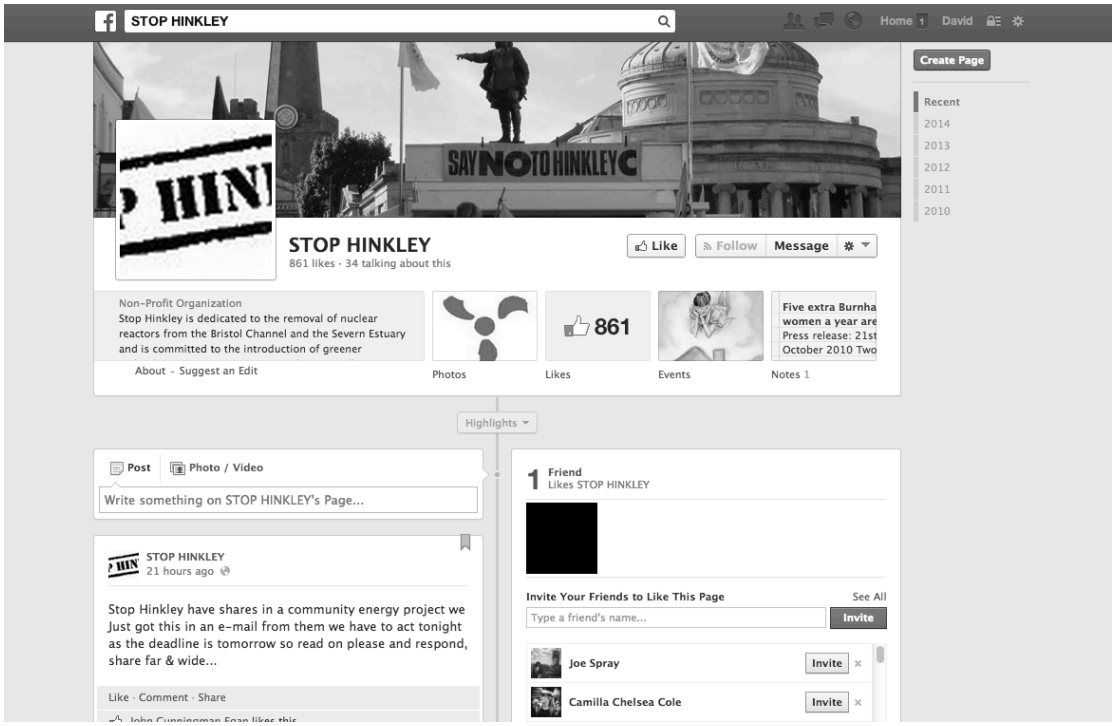


Image 9. Stop Hinkley Facebook Page

Stop Hinkley

To show that these digital demonstrations are not an isolated occurrence, I will now turn to the page of a related organization. Stop Hinkley is a regional anti-nuclear group which predates Facebook by about 20 years. Stop Hinkley was originally Stop Hinkley Expansion (SHE) and participated in the 1988-9 public hearing into Hinkley C (Aubrey, 1991). Stop Hinkley's Facebook page has technically more likes than Stop New Nuclear (861) and more overall comments despite being a smaller, more focused group. There are however less frequent posts on the Stop Hinkley page.

Using the same process, I will create another tri-partite Graph for the year 2013 when another set of anniversaries and protests were planned.

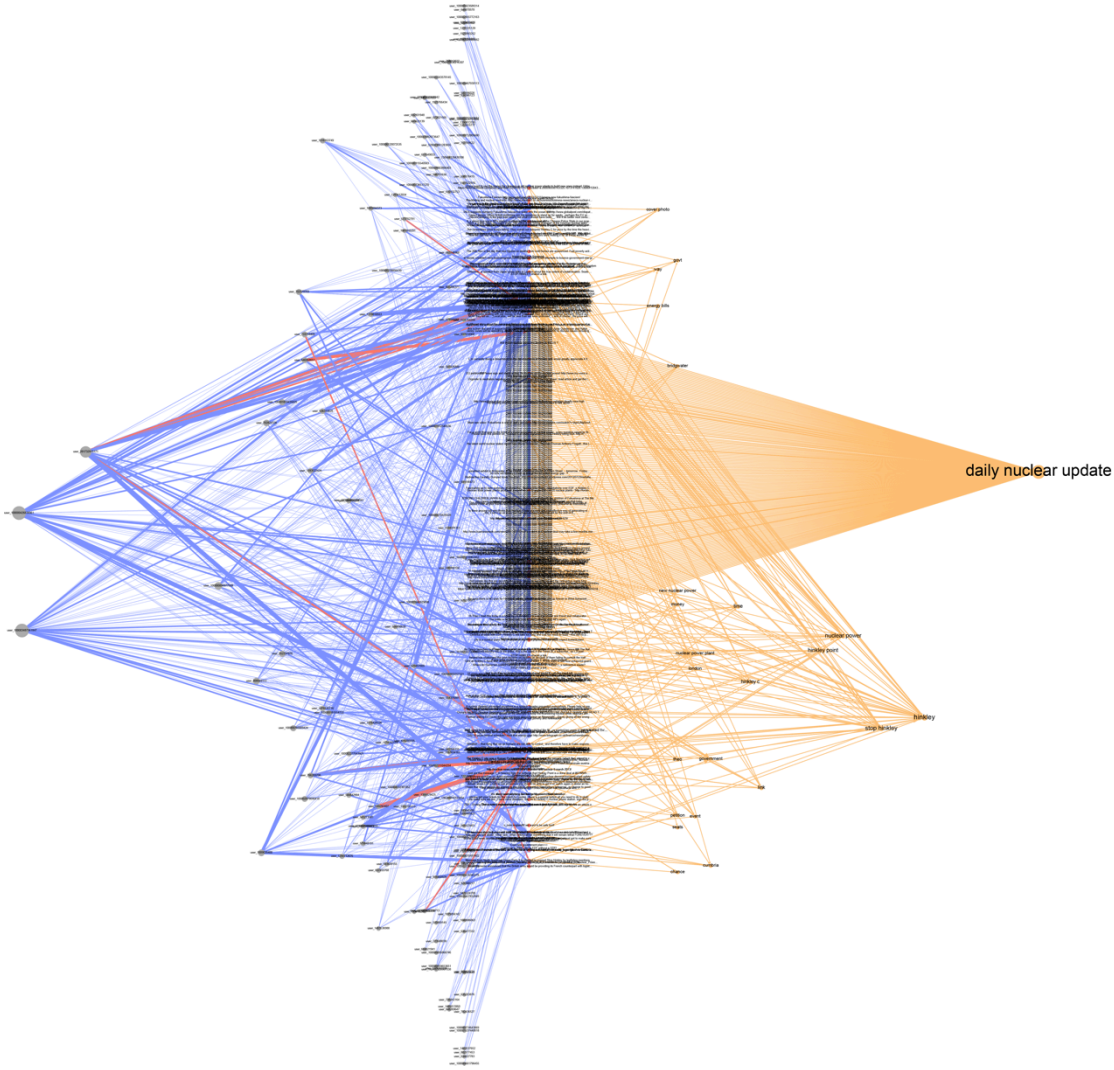
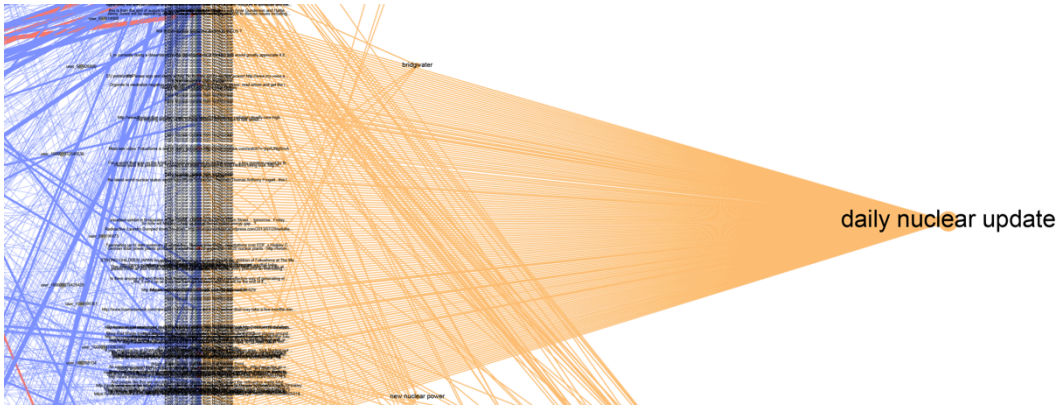


Figure 12. Stop Hinkley Tri-partite https://goo.gl/Jv5dCP: Each graph is composed of three columns (users on the left, posts in the middle, words on the right). The posts are coded red for page or admin-generated posts and blue for user-generated posts, this is also true for the lines connecting users to posts.

The main difference between this graph and Stop New Nuclear and the email list is that while there are still references to offline protests there are far more references to news stories and digital demonstrations, increasingly over time, something which would not be spotted with the two separate network graphs in static slices.



Detail of Figure 12.

At the beginning of the year, the group members tended to use Facebook more like a traditional email list, to organise offline events and establish ties with other groups: e.g.

> Message from North Somerset Green Party It is planned to support Stop Hinkley by leafleting members of the public_ and engaging attendees at the Your North Somerset event at the the Winter Gardens_ Weston this Sunday 27th. ...

It was also used for sharing information: for example the noticeable sequence of repeated text in the middle, designating a series of posts titled Daily Nuclear Update in the detail above. These were posted by No2Nuclear, cross-posting updates from their blog, again, revealing the interconnections between platforms. As the left of the graph reveals, this generated very little engagement, at least in Facebook's terms.

But as the year went on they started to deploy the page more in the service of 'digital demonstrations'. This primarily consists of e-petitions, at least 7 in the year 2012-13 some of which are from You.Gov, but they also hi-jack other platforms:

> Leave your own comment on the Guardian thread:
> anyone fancy having a chat with EDF fans? http://youtu.be/NwTeraRtd6g

Posts like this are encouraging Stop Hinkley users to intervene in a particular article, comments or video comments because they are potentially being overrun by pro-nuclear voices. They are balancing opinion where it is perceived that pro-nuclear supporters are *overrepresented*.

While the goal of the e-petition is relatively self-evident: (petitions with 100,000 signatures will be discussed in parliament) the other interventions such as arguing on YouTube or appearing in a guardian thread are less clear.[154] Through regular posting, liking and sharing and interventions on other platforms, they are trying to achieve a certain market-share of people's online attention.

The tri-partite diagram gives some indication of how Facebook draws users into this logic of metrics and analytics. Because ANTA has not fully cleaned the terms, there are some interesting accidents such as Facebook specific key words like 'share' and 'cover photo'.[155] These increase towards the top, the most recent posts around Nov 2013. While the members originally deployed the page much like an email list they become increasingly savvy in the Facebook visibility game. This is evidenced by the following post by the page admin:

> Please folks SHARE the news that we put up here, I know a lot of you are reading it but we need you to share it! Please do give us feedback about why you don't share it if you don't. **We are all the media now** - the corporate media cover-up the nuclear issues but social media such as Facebook twitter etc. are our only hope of reaching the masses - but in order for that to work we need to extend our reach, this means we need you to SHARE our stories - go on be brave put our stories on your wall so that your friends can see them we need to see more of he shared via Stop Hinkley on peoples walls. So go on get clicking!
> *emphasis added*

So some members of the activist groups explicitly see the role of Facebook as an alternative new medium, not just as an organising tool, but this means at least ambivalently adopting Facebook's definition of the conditions for success, the wider spread of their content, likes shares and comments. This is however not the only way that Facebook helps define legitimate participation.

Facebook's Version of Participation

[154] One of my key informants (an admin on another page) claimed the goal was to '…get the message out and recruit new followers'.
[155] There are also some posts which either contained no analysable text (a link share with no comment) or no text shared with other posts. This may indicate that ANTA could be more inclusive with what counts as an entity in this case, or ideally the same method could be applied to the text of comments as well as posts.

Sometimes users complain about astroturfing on other platforms, but occasionally Facebook's own platform politics enters the picture:

> ok dodgy goings on with FB as ever there is anews story on itv west that I want to share here except I can_t share it as stop hinkley so i shared it to nikki_s page and she tried to share it here twice but - well - i can_t see it here can you? it_s about richard cotterell former euro mp... it should be in the box that says recent post by others but can_t see it there either hmmmm....b**** FB!
>
> some person is checking your mail. very dodgy.

It is a common occurrence for posts to be automatically removed by Facebook either through an algorithm which tries to find, for example, nude pictures or posts which incite violence. Protests of course may be confused with the latter – though it is unclear if this is deliberate censorship or accidents of a computer programme.

While users may be debating what counts as legitimate participation on social media, Facebook, through its design and coding, imposes its own ideas of 'proper participation'. Bucher for example notes a hierarchy of actions which favours friends who use the chat function or videos as opposed to text, but the more pernicious omission above frames political participation as, implicitly, *peaceful*.

Facebook also inscribes assumptions about the proper participant. One of my informants expressed concern that her account was shut down because she did not use her real name. In older technologies of the web (email lists included) the use of usernames and avatars was encouraged to protect identities, but this is actively discouraged by Facebook (Morozov, 2012).[156]

To really draw out some of the politics of Facebook, the way that the design of the platform formats participation, it is helpful to compare this activist usage of Facebook to that of a very different kind of group, the French energy company EDF, who will manage the future Hinkley Point plant.

[156] Available from: http://www.michaelzimmer.org/2010/05/14/facebooks-zuckerberg-having-two-identities-for-yourself-is-an-example-of-a-lack-of-integrity/ (Accessed 28 September 2015)

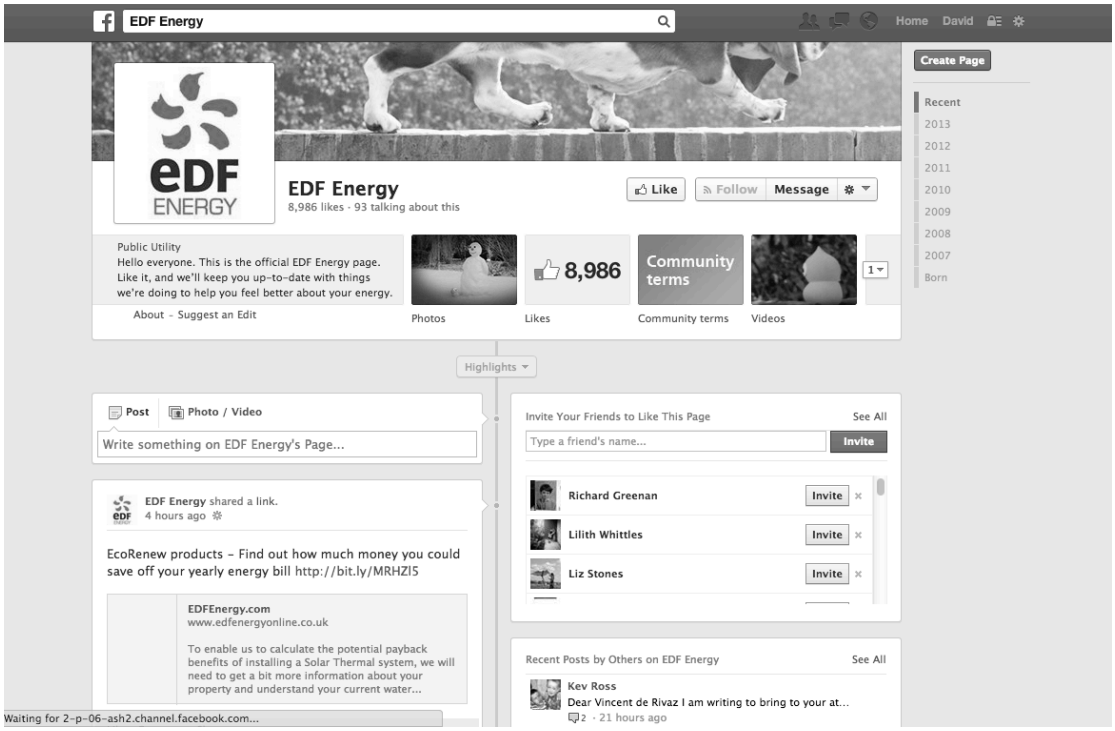


Image 10. EDF Facebook Page

While Stop Hinkley had nearly 900 followers, EDF has 9,108, more than 10 times as many likes. This is what I mean by activists not being able to compete, if participation is understood in terms of the quantity of likes.

Firstly, one of the main differences between the corporate EDF page and the activist page is a 'community guidelines'. These include, importantly, the right to delete posts that…

> 'defame, abuse, threaten, incite violence against, or otherwise violate the legal rights (such as privacy and publicity) of other users, the moderator, Us, Our affiliates (including all companies within the EDF Group), or other partners or affiliates of EDF Energy.'

So one methodological challenge is that posts which have been removed can not always be retrieved. Removing posts of course has its downsides because, as we saw on the Stop Hinkley Page, users can complain loudly about posts disappearing.

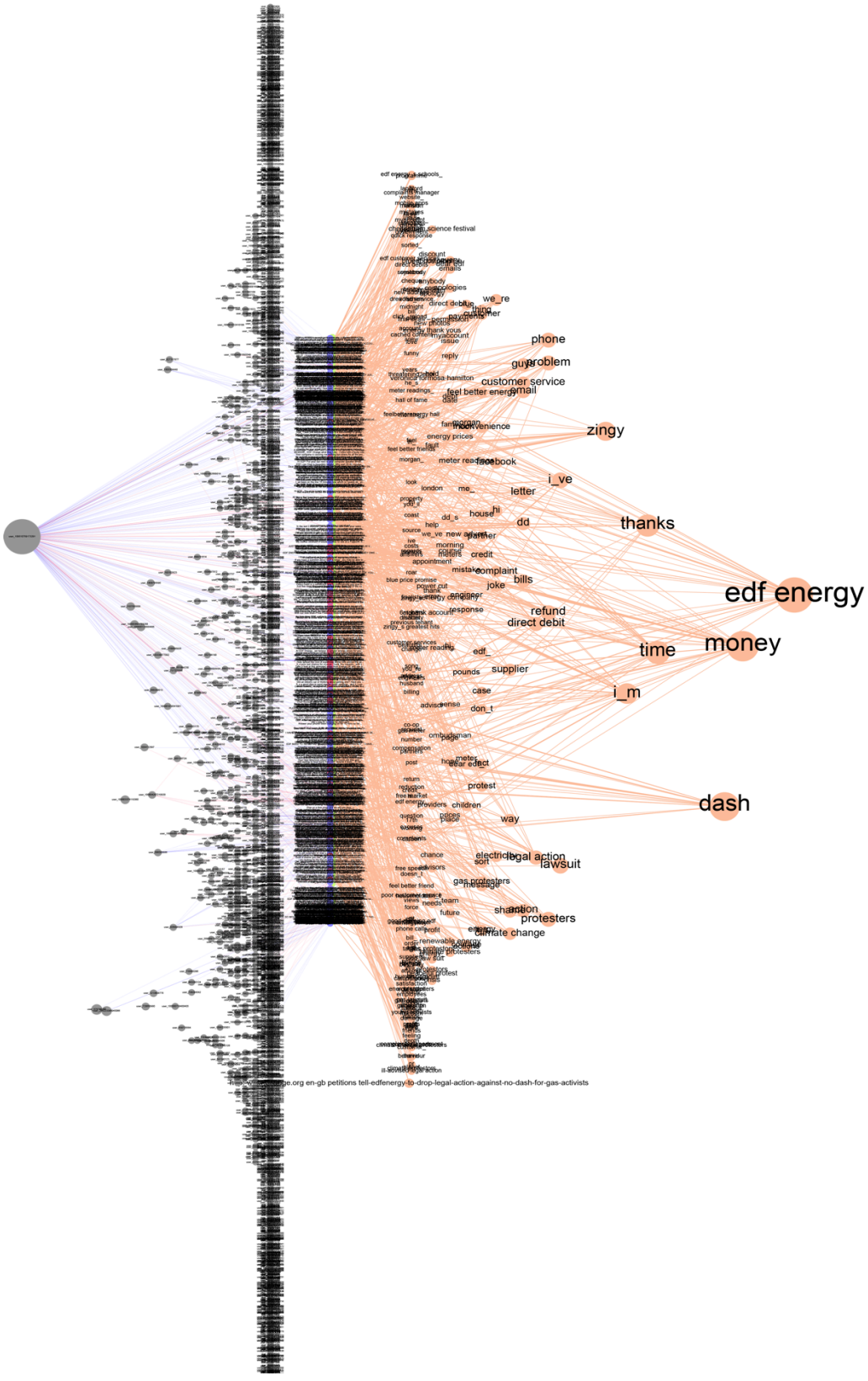


Figure 13. EDF Tri-partite https://goo.gl/35zij2: Each graph is composed of three columns (users on the left, posts in the middle, words on the right). The posts are coded red for page or admin-generated posts and blue for user-generated posts, this is also true for the lines connecting users to posts.

Looking at the tri-partite diagram, in contrast to the page as it appears on the screen, the EDF page appears to be mostly a torrent of customer service related abuse. The terms on the right include 'help', 'problem', 'fault', 'house', 'letter', 'mistake', 'refund', 'joke' etc.

A typical example is:

> dear cowboys_ I changed supplier in September_ why have you billed me to 17/10?
> 2013-10-23T17:03:38+0000

Followed by one of many semi-scripted, 'boilerplate' answers by the page admin:

> Hi Lee, I am very sorry to hear about this. Please email details, including your account number, full address and contact information to socialmedia@edfenergy.com ...

But these complaints do not appear of the front page, they are buried in the 'Posts by Others' section (see above), which is much harder to read and access. Facebook thus allows the page admin to manage visibility. If however the admin engages with the user in a more active way, the posts may be picked up by the algorithm and appear on people's news feeds, so the best strategy is to respond once and ignore.

However, Facebook is not just a site optimized for corporate control, it can also be a site of creative resistance. At the bottom end of the graph there is a cluster of activity with some decidedly non-customer service related words like 'Dash', 'legal action' 'protest': referring to EDFs lawsuit against protesters from an environmental group called No Dash for Gas. On the run up to the verdict, activists bombarded the EDF page with abuse and slogans.

Negative posts can always be deleted, so the more effective strategy for the dissenters may be to hi-jack a post by the page admin, which is automatically more visible. Many of the recent page-generated posts deal with Ziggy, the non-threatening anthropomorphic pilot light (see Image 10):

> Zingy and Morgan_s ad will be in Scott & Bailey tonight. Two great duos... one solves crime_ and the other brings you Feel Better Energy!
> 2013-04-17T20:00:01+0000

0 likes

EDF-off and take your toxic nuclear slime with you

I really don't care about Zingy and the advert. Why is Facebook forcing this ridiculous EDF advertising onto my Newsfeed?

Maybe get Zingy & Morgan to visit the elderly and see how they struggle supporting your huge profits.

Note that EDF regularly use the Facebook paid-for-feature of promoting posts in people's news feed, but everyday users, whether they are actively protesting nuclear power or not, can then hi-jack these widely circulated posts to their own ends. This example of a digital demonstration is an important gesture, which strangely calls to mind graffiti, it serves a purpose of communicating something, but perhaps more importantly it proclaims that a coherent and articulate opposition exists.

05. CONCLUSION

This chapter has argued that we need to bring traditionally quantitative and qualitative traces closer together to qualify and contextualise metrics but also scale up the act of qualitative analysis and get a better sense of trends over time. This is important because while platform metrics do not capture the complexity of activities online, these metrics and the visibility games that come with them format and drive activities on the platforms and cannot merely be brushed aside.

I arrived at this through considering the contribution of online platforms to the controversy over Hinkley Point, a skirmish within the wider issue of nuclear power in the UK. While Facebook is still used to organise events on the ground, another potential contribution to controversies might be in terms of online only interventions which either advance knowledge and articulations of the controversy or merely demonstrate the existence of a significant opposition, perhaps through an accumulation of likes but also advancing alternative articulations in the media or even symbolic interventions. The concept of 'digital demonstrations' allows for this more issue-based reading of the contribution of platforms.

This leads to the second point, that Facebook itself encourages evaluating these interventions in quantitative terms, numbers of likes, shares etc. This is in sharp contrast to email lists which have no such analytics, or other offline activities, where activists are cautious as to how seriously they can measure impacts – like the numbers of attendees. This goes back to Mica White's original warning about clicktivism. Firstly, this quantifiable, marketing logic eclipses other ways of understanding activist interventions which are more symbolic and affective. Secondly, this is a game that activists cannot hope to win against better resourced mainstream media outlets and corporate actors because of asymmetries scripted into the design of social media platforms. While I have attempted to complicate the assumption that online actions are a poor replacement for offline protest, it is important that users but also researchers do not uncritically adopt the device perspective of Facebook (and other social media) which reduces complex activities to calculable, analysable traces.

As with the other chapters, activity measures such as the frequency of posts or likes are decent indicators of controversy, as a gathering of actors around an object, but this must be confirmed by investigating the content of the activity which one can get a sense of by consulting the left side of the graph. In contrast to the previous chapter where digital traces were compartmentalised, Facebook's formatting allows text and number to be analysed together. Although I have been talking about Facebook and 'likes', the same could be said for other online platforms and practices including the other venues that activists travel to: Youtube comments have their own ranking system of up-voting and down-voting; news comments such as those on the Guardian website have a similar 'top comments' section decided by algorithm; e-petitions of course have their own quantitative logic.

As I suggested, platforms like Facebook create a game of visibility, but the game is sometimes rigged. Design features and Facebook algorithms, police what kinds of actors and what kinds of content are admissible on Facebook and it facilitates the suppression of dissenting messages. However, this is never given and can always be contested by creative activists. However, if activists use metrics such as 'likes' and 'shares' to monitor these activities, these metrics actually tell us very little about how their information spreads, other than an expansion of the potential audience. We do not know how information spreads through these networks. This problem will be to some extent taken up in the next chapter.

VI. NO AMPLIFICATION WITHOUT MODULATION: INFORMATION DIFFUSION ON TWITTER

Early on in this thesis, I lamented the fact the most ethnographically inclined studies of controversies did not address the contributions of media in any great depth. I speculated that this was on one hand due to a suspicion that the media only circulated and transmitted information from science, not adding much to the proceedings (Pinch, 1994) except perhaps distortion (Hilgartner, 1990) but also because STS only had access to models of information flow inherited from media which were based on sender-receiver models, circuits or media specific analyses (Lewenstein, 1995b). These are of course necessary reductions of a complex phenomena because, in the case of broadcast media, it is nearly impossible to convincingly track information flows between producers and audiences.[157] However, social media platforms, with their extensive documentation make the study of information diffusion more feasible, through it is important not to overstate this point.

In the last chapter, I examined the process by which activist groups try to publicise articulations of a controversy and themselves, partly at least through the accumulation of likes and shares which trigger Facebook's algorithm to distribute content to friends and thus potentially friends of friends and so on. But due to Facebook's privacy settings and proprietary algorithms this process remained mysterious, summed up only in quantitative terms. In this chapter I will attempt to get into the mechanics of information diffusion.

I will approach this problem, yet again, through a particular controversy, or rather set of controversies. First there was the official announcement in March 2013 of planning permission for Hinkley C plant, alluded to in the last chapter, which is the first of a new generation of nuclear power plants in the UK. This elicited condemnations from the anti-nuclear community but made little impact elsewhere. Secondly there was a blackout at Fukushima which threatened the stability of the containment operation.

[157] It is arguably easier to track information flows in science given the smaller, inward looking audience and the stable formatting of scientific articles over a long period of time.

This was not covered widely by the mainstream press and for concerned actors in certain corners of the web, this suggested the media was ignoring or supressing the story. But despite the relative lack of coverage of either of these controversial events, one venue where these controversies were discussed and disseminated was the microblogging platform Twitter.

Founded in 2006, Twitter is one of the newest so-called 'social media' platforms. Twitter operates much like Facebook (it is based on Facebook's status update; van Dijck, 2013) except that it is more 'public' – users can follow the posts of almost anyone – with the exception of 'private' users. Yet just because content on Twitter is more or less 'public' does not mean everything is equally accessible. Just as with Facebook, Twitter creates another visibility game, this time based around making certain content 'trend' – rising to the front page of Twitter. So one potential contribution of online platforms like Twitter is to promote or disseminate information about controversies in lieu of mainstream media coverage. Yet another contribution has to do with articulations of the issue: is the Hinkley plant presented in economic terms or health and safety terms and which other events and controversies does it become attached to.

The argument of this chapter, in contrast to conceptions of media or particularly social media as merely amplifying or hyping content, is that any circulation necessarily involves changes in the content of articulations, and these changes have implications for further circulation. This requires looking at the reciprocal relationship between content and infrastructure – what is said about the controversy and the socio-technical machinery of platforms. This will be intuitively accepted by most STS scholars but the phenomena of information diffusion, again, falls in the gap between different sorts of data and different methodological techniques somewhat encouraged by social media data structures. Twitter however makes possible the reconciliation of content and infrastructure but this again requires going somewhat against the devices and objects foregrounded by platforms and their APIs.

In what follows I will consider different mechanisms for how content can spread from one user to another, pointing out that most of the existing analysis techniques assume that we know in advance the vehicle(s) of diffusion. I settle on tracing hyperlinks, firstly because this is a stable object through which modes of diffusion can be monitored and secondly because it opens connections to the world outside of Twitter as a platform, particularly online news and blogs. This again requires going against the

most obvious devices Twitter offers up for analysis: such as hashtags or mentions of users. I will then propose an innovative way of describing the trajectory of these links using colour coded strips which I will apply to top news stories related to the controversies above.

However, first I again need to consider on what terms to understand the contribution of Twitter, because according to the literature it may be productive of exactly the sort of hype or banal media effects which might rightfully be dismissed by controversy analysts in the past.

01. TWITTER AND CONTROVERSIES

Despite the deep integration of Twitter into journalism and political life (Bruns and Burgess, 2012), it is by no means self-evident that Twitter is an appropriate platform through which to study the unfolding of socio-technical controversies. While Wikipedia allows users to potentially contest and parse available expert accounts, and Facebook can organise protests in offline and online space, it is not immediately clear, based on the past literature, how Twitter may contribute.

Banal Content, Rumour and Spam

Firstly there are longstanding concerns about the trivial content of Twitter messages. Richard Rogers (Rogers, 2013a) explains in a recent article, since its founding in 2006, Twitter have gone through several phases, as has the academic literature on it. The first incarnation of Twitter research focused on the sharing of banal personal information: 'what people are having for breakfast'. Vincent Miller (2008) called this 'ambient intimacy' the implication being that the substance of what is said on Twitter is less important than the connections made and sustained. While controversy analysts may be interested in the rise and fall of networks on Twitter seem to be defined specifically by the *everyday* rather than the controversial.

However, phase two of the research recast Twitter as more of an event-based news medium. In volatile events such as natural or man made disasters, (Doan et al., 2012; Murthy and Longwell, 2013; Vieweg et al., 2010) riots (Procter et al., 2013; Vis, 2013b) or revolutions (Meraz and Papacharissi, 2013) Twitter enables a form of on the ground

citizen journalism which by-passes the relevance defining role of traditional journalists (Hermida, 2010) as already discussed in Chapter II and IV (Murthy and Longwell, 2013). So there is a potential that, in the case of rapidly unfolding controversies, non-experts and non-journalists can advance claims or facts about what is happening.

But still there are doubts about the salience of these claims. Many studies of disasters (Castillo et al., 2011; Mendoza et al., 2010) view Twitter in relation to the spread of *rumour* or hearsay. However some scholars see Twitter as a rumour-quashing engine in which falsities are quickly corrected. As Rogers notes in the case of the London riots, when users started claiming that the rioters had released animals from the London Zoo, these pranks were quickly weeded out by the larger Twitter community (see in particular Procter et al., 2011). So even if Twitter facilitates the advancing of alternative facts and claims, it is unclear if the claims themselves are worthwhile or have much of a shelf-life (see Bruns, 2012) before they are either denied, forgotten or assimilated into so called mainstream media accounts.

Twitter is now, according to Rogers, in its third phase, in which it is analysed regularly as an archived object, by both social scientists, corporations and government, often as a way of gauging 'public opinion' or consumer desires. It has in this sense also been used by STS researchers to study the dynamics of issue-formation, before institutional actors arrive on the scene (Marres and Moats, 2015; Marres and Weltevrede, 2013) through with much more consideration of hype and publicity effects. Twitter today, as I will show, is also heavily marked by the emergence of spam and bots: or non-human users who tweet automatically. One old estimate states that bots are responsible for as much as 20% of all content.[158] Whether spam, rumour or banal content, the 140 character limit and the culture of Twitter seems to put a limit on the potential of Twitter as a platform for intervening decisively in controversies.

The Hinkley C Announcement and Fukushima Blackout

Again, as with the other cases in this thesis, the most appropriate terms on which to approach platforms only becomes clear through particular controversies. Among the other platforms I was monitoring, I had been collecting Tweets related to the keywords (Fukushima, nuclear, nuke, Hinkley, EDF) using the Digital Methods Initiative's TCAT

[158] Available from http://variety.com/2013/digital/news/twitters-spam-headache-more-than-10-mil-accounts-might-be-bogus-1200694134/ (Accessed 9 September 2015)

package (Borra and Rieder, 2014). Within this keyword defined space, I was looking for bursts of activity (liveness) around anti-nuclear protests but instead, the greatest spikes (of both Hinkley and Fukushima) concerned the a series of events around March 19-20th. Although these provided starting points, a rough index of controversiality, as in the other chapters it is important to move past frequency measures to determine which kinds of activities these bursts represent – are these important interventions in a controversy or media-specific effects which may be less important to the cartographers of controversies.[159]

Just over a week after the second anniversary of Fukushima, which prompted worldwide protests and (social) media commentary, on March 19th the UK government granted planning permission to French energy company EDF to build the first nuclear power plant in the UK in nearly 20 years. Although the future of the plant still remains uncertain at the time of writing, this was a major milestone in the history of UK nuclear policy. But with cosmic irony, while the press releases were being readied and speeches prepared, Fukushima re-emerged in the news. First on the 18th, TEPCO announced that it had identified a fish with record levels of radioactivity in its nets. Second, and more worryingly, power was mysteriously lost to reactors 1, 2 and 4 of the crippled facility. Constant power is needed to maintain the water pumps, which cool the cores and prevent further meltdowns or pressurised releases of radiation. A few days later on the 20th it was revealed that the power cut was likely due to a rat, which had been fried while chewing electrical cables.

As Andrew Barry notes, a knowledge controversy must be '…conceived of in its relations to a moving field of other controversies, conflicts and events, including those that have occurred in the past and that might occur in the future.' (Barry, 2012: 330). He gives the name 'political situations' to this process of making (or denying connections) between a broad range of issues, politicizing or depoliticizing them in the process. This term is helpful for understanding what is at stake in connecting or not connecting these distributed controversies.

It is also important to note that that the Hinkley C announcement was largely articulated in a de-politicized and technical manner. The announcement of Hinkley C

[159] It will also become more clear in the next section why keyword defined metrics are only one of several ways of defining a particular controversy on Twitter.

was an example of what media sociologists sometimes call 'pseudo events' or carefully orchestrated happenings purely intended to attract the attention of journalists (Boorstin, 1961; Tuchman, 1978). Some of the on-the-ground protests I described in the last chapter might function in this media baiting way, but most protests have an open-ended character to them, where pseudo events like press conferences are intended to *control* possible outcomes and interpretations. The Hinkley announcement was largely presented in what Peoples (2014) calls an 'energy security' frame – presenting nuclear as a pragmatic choice in light of limited national supplies of fossil fuels and the impending threat of climate change. This frame marginalises concerns about nuclear as a health and safety risk, which is how the Fukushima events were articulated. In Chapter II, I pointed out that while frames deployed in this way might presume too much fixity in the meanings and too much focus of discursive practices at the expense of technological contributions, they do speak to important capacities of online media to articulate controversies or issues *differently* than the mass media. In this chapter I want to add some contingency, materiality and process onto the bones of these media frames and firstly this means understanding better how frames are distributed.

Between Content and Infrastructure

So perhaps what is distinctive about Twitter in these sorts of media events is its capacity to *distribute* external content and (re)frame it. This problem takes us further away from classic public science controversies over knowledge claims to confront questions more specific to media. In his study of a sample of French Twitter users, Bernhard Rieder (2012) suggests that the advancing of claims or facts, may actually be quite rare, and specific to disasters. Instead he argues that Twitter users are more likely to add a bit of 'spin' or 'twist' to *external* content using hashtags or discursive commentary, which he calls 'refraction'. This 'refraction' could be important way of describing how social media users can contest mainstream media articulations or narratives of controversial issues, on an everyday basis.

What Rieder is challenging with this observation is what he sees as the dominant paradigm of current Twitter research: 'information diffusion'. This constitutes a range of approaches from cultural memetics[160] (Blackmore, 2000) to theories of contagion

[160] 'Memetics', builds on Richard Dawkin's concept of the 'meme', modelled after the biological gene which is transmitted though successful procreation, but applied to the study of culture. In

drawing on the work of Gabriel Tarde (Kullenberg and Palmaas, 2009). There is not space to discuss this broad literature here, but what I want to highlight is that these approaches often, but not always, associate the spread of information with networks (in particular of the digital variety).[161] The key problem for Rieder, however, is that these approaches tend to separate out the infrastructure or medium through which information spreads and the information or content itself.

Now, sharing could be studied in a variety of ways. It is possible to look at the proportion of coverage quantitatively, the equivalent in media studies of looking at 'column inches' devoted to a topic. Murthy and Longwell for example find that, even in the case of disasters, Twitter users share far more links to mainstream media sources than to alternative ones, which raises doubts about the extent to which Twitter can offer so called alternative messages (2013). In a more media-specific way, one could study Twitter's metrics of 'trending' – which content is picked up by various algorithms as being popular that day. It is crucial to understand the reflexivity of actors who attempt to game these algorithms (Gillespie, 2014) but studying the process of trending empirically is difficult because the algorithms themselves are proprietary and they speak to what is, crudely, popular in terms of volume, which might direct the researcher towards advertising, spam and celebrity content. Also, such an approach places the focus on Twitter itself, rather than how Twitter is involved in the articulation of specific issues.

There is also a tendency when focusing on the quantitative volume or popularity of content to smooth out the divergent ways that information spreads, attributing explanatory power to algorithmic logics or network structure. I will propose, following Rieder, to study the dynamics of *how* content travels: what devices, technologies or resources are employed and which of these strategies seem to be the most effective in

---

common parlance, the 'meme' has come to mean mass repetitions of jokes, images or concepts in platforms like Youtube, Reddit and Twitter.

[161] Although Gabriel Tarde's work on imitation and innovation has inspired strands of this social theory, more sensitive readings of his work, such as Sampson's (2012) would in contrast show spreading necessitates changes in that which spreads; no two imitations are the same. As alluded to earlier: while for Tarde, beliefs and desires are transmitted, individual instantiations of them in acts of imitation are unique due to individual sensations. Recently, Tarde's work has been commonly associated with networks largely due to the work of Bruno Latour (Latour, 2010; Latour et al., 2012). But it is important to remember that Tarde's idea of society is always emergent, that is, networks are not merely a cause but a consequence of contagion-events. Although Tarde's contagion, which is a psychological process, is difficult to study empirically, the approach proposed in this chapter could be seen as building on Tardean statistics, which map variations in repeated phenomena over time (Barry, 2010; Didier, 2010).

causing content to be taken up by others – is it the discursive content, the follower networks or the deployment of bots and scripts? I will not, however, be able to answer this question definitively based on this small case, but in the next section I will briefly explain several different mechanisms through which Tweets can be generated and made available to other users.

02. TECHNOLOGIES OF DIFFUSION

It is important to appreciate the diversity of practices on Twitter, which are deployed differently by different types of users at different times, because scraping and analysing Twitter often necessitates privileging certain behaviours and digital traces at the expense of others (Marres and Weltevrede, 2013). Both quantitative and qualitative researchers must ultimately circumscribe their data by, for example, a networked group of users; a hashtag or keywords; identical retweets etc. When this is performed uncritically it can present particular practices and features of Twitter as central explanations of information spread, when this centrality is precisely what needs to be explained. This is yet another way that the particular affordances and data structures of a platforms directs how it is studied. In this section I will, using recent literature about Twitter, detail several of the interlocking devices through which information might pass from one user to another before attending to how information spreads 'in action'.

Networks

One of the key ways users can receive information is to 'follow' the tweets of other users, so that their messages will show up in their 'feed' (a stream of incoming Tweets). This networked way of receiving information makes Twitter similar to sites like Facebook, with the key difference that following need not be reciprocated: it can be asymmetrical. But even more so than Facebook, it is important to be critical about the status of these associations between users.

The number of followers, much like the number of friends on Facebook, likes or counts of unique users, represents a metric which can be leveraged for advertising revenue or financial gain, much like television viewing figures. So it is important to realise that following can and will be 'gamed' for commercial / personal advantage: for example many inactive accounts in follower networks may be 'fake' users which can be

purchased by the thousands to boost followers (Vis, 2013a). There are also a range of potential uses for following from friendship to 'subscribing to' even 'monitoring' or 'stalking.' Currently however, the Twitter API does not allow for the possibility of visualising these ever shifting networks of users and followers.

A related way of studying information diffusion on Twitter is through 'mentions'. Twitter has over time developed several Tweeting conventions, most notably the use of an @ symbol to denote a user – e.g. @davidjmoats. Whenever a users 'mentions' another user in this way, the recipient is notified of this through the interface. Mentions can take many forms, from a conversational question or a prompt in order to illicit a response, a tacit thank you or a show of appreciation:

> @roilogolez Sir your info was wrong yesterday. China has 28 nuclear reactors under active construction not just one. [URL]

The above Tweet is engaging the user @rollogolez in conversation, hailing them effectively, but there are other uses for @ mentions such as giving credit, in the case of a 'retweet'. Retweeting is when all or part of a Tweet is reproduced and credit is given to the originator of a message:

> RT @HuffingtonPost: Nuclear power: Damned if you do damned if you don't?

Most frequently this is written as 'RT @username' but also 'rt @' or 'retweet @' followed by the contents of the original tweet, often truncated to accommodate the extra characters needed for the user name '…'. 'Via @username' often denotes that the content has been paraphrased. Users may also place the RT at the end of the Tweet or combine a retweet with a mention to share with users who might not already have seen the content. dannah boyd et al point out that retweets are also a strategic way of alerting another user to one's presence and gaining a potential follower (boyd et al., 2010).[162] From now on I will refer to uses of @ generally as '@ mentions'. What @ mentions do is to alter what Murthy refers to as the 'participation framework'.[163] This

[162] These textual conventions have become so popular that a 'Retweet' and 'Mention' button have been added to the Twitter interface to, for example, automatically retweet a particular tweet in the format 'RT @username original message.'

[163] Murthy is referencing Karen Knorr-Cetina (2009) who attempts to update Goffman's interactionist 'situation', a face-to-face interaction in which two or more parties are physically present, to mediated interactions, such as between two stock traders, by adding the word 'synthetic'. In synthetic situations, parties may be present in time but not in location although

draws on the micro-sociological theory of Goffman (1981) and describes how utterances (in this case Tweets) imply a particular perceptual range – who can take in the utterance. Those within the perceptual range have a 'participation status' relative to the utterance – are they commanded or invited to respond or merely to listen?[164] So sharing not only happens through a network, it also may grow the network.

@ mentions are often visualised as directed networks of users connected by mentioning each other, as a way of representing the flow of information. This is based on the assumption that users will normally acknowledge the original source of a tweet, making visible the routes through which information flows. Meraz and Papacharissi (2013), who use Twitter to study the Egyptian revolution, assume from existing literature on social networks that the most 'mentioned' accounts (the highest in-degree count) will be the most important in driving information flows. So they reduce the dataset to a corpus of users with the highest @ mentions. Through formal properties of networks they make claims about the centrality of certain users in information flows. But this approach automatically excludes the contributions of users who chose to not acknowledge their sources or who receive information in different ways. If one scrapes for conversational elements like @, then one gets *only* conversations.[165]

So there are some limitations to using networks of @ mentions to study information diffusion: firstly @ mentions have many uses other than simple attribution of where content originated; secondly, @ mentions do not necessarily map the network through which content spread but may reveal a network being built *as a consequence* of the content spreading; thirdly, there are other ways that content can spread which do not leave visible traces such as @ mentions.[166] In what follows I will argue that shifting the participation framework can occur in less obvious or traceable ways.

---

sometimes interactions can be deferred over a long space of time. They are also mediated by and equipped with scoping 'technology'. But as I will argue due to bots, perhaps the synthetic situation is not synthetic enough! Why does it make sense to start with human-human interactions when so many interactions are bot-bot?

[164] So if user A produces a message and another user B retweets it, prefaced with an 'RT@user_A…' a third user C may retweet but attribute the message to @user_B meaning the source is disguised. Users cannot see the whole chain of retweets, only people they follow, which means messages can transform over time through these friendship networks.

[165] Add to this that @ mentions can denote a variety of behaviors (Boyd et al., 2010). An @ can be used to *attribute* content to someone or *solicit* a response (@ mentions register as notifications on the user's interface). Some users retweet only the user they received the tweet from while others acknowledge the originator of the message or the full chain of users.

[166] This is not to say that @ networks are invalid – they are highly appropriate for analysing modes of sharing in which making / maintaining social connections is important. But there are

Bots and RSS

In recent years, Twitter has become completely inundated with bots and scripts which tweet automatically based on certain triggers or criteria.[167] The most common are what I will call 'forwarding services'. These include websites and apps like Twitterfeed, dlvr.it, IFTTT and Hootsuite which are based on RSS technology and are set up to Tweet a message whenever an article on a website is published or updated, whether or not the owner of the account is even awake.[168]

Users of Twitterfeed (twitterfeed.com) for example can link up to highly specific feeds based on metatags for the article category (business, entertainment, technology etc) and customise their tweet with a personal message including hashtags or @ mentions to tailor it to these feeds. Other services like IFTTT (If This Then That: ifttt.com) can also be triggered by events on Facebook or LinkedIn and a programmer could design a bot to Tweet a message based on what is 'trending' that day.[169] The point being that tweets may arise from completely backchannel sources like RSS, which may themselves generate @ mentions or illicit followers but do not originate *as a result of* direct mentions or following someone.[170]

However, Wilkie, Michael and Plummer-Fernandez (2014) make the key point that distinguishing between human and bot is difficult because, while humans may set up robots to do their bidding, other technologies, such as the semi-automated Tweet Button, embedded underneath many news articles, prompt human users to act very

---

still other ways in which information can travel which are not captured through these *traceable* interactions.

[167] An old website called Bot or Not (Available from http://web.archive.org/web/20130430025727/http://botornot.net/project Accessed 29 September 2015) estimates that out of a sample of 18,000 accounts 15% are bots, another 16% are 'probably bots', while 33% are human and 36% are 'probably human'. What is important here is not the proportion of human / bot, which will of course vary wildly in different contexts, but the fact that Bot or Not must draw the distinction as a spectrum!

[168] RSS or Really Simple Syndication is a web protocol that allows users to receive notifications when particular pages are updated or when new pages are added to websites, often including the Title' and a short Description of the content.

[169] For example, the following tweet is based on what is trending' Trending: Hinkley nuclear plant awaits go-ahead [URL] Rippla News 5:29:02 AM

[170] I have confined this discussion to various online means of information spreading but of course there are offline ways: User A in physical proximity to User B tells her 'Hey, did you see my Tweet?'

bot-like.[171] The authors propose thinking of Twitter as a socio-technical assemblage of human and non-human actors rather than artificially distinguishing between the two.

Hashtags

Another popular means through which users can receive and share information is through hashtags, which are a popular means of data reduction because they are topic specific and user-defined rather than researcher-defined unit of analysis. Hashtags are when a # is placed in front of a word or phrase with no spaces, for example: '#Fukushima'. If enough users employ a hashtag it can begin to 'trend' which means it will feature on Twitter's front page for the user's chosen region and be picked up by various algorithms and other platforms monitoring Twitter. Users can use the search function of Twitter to 'tune into' a popular hashtags much like a radio station or as Diraj Murthy puts it, like internet chatrooms where interested users discuss topics under user-generated headings (2013).

But there are some problems with scraping by hashtags in terms of information diffusion. Firstly, and quite obviously, it is difficult to draw conclusions about the centrality of certain hashtags over others when one starts with a hashtag to define the data set (Tufekci, 2013). When hashtags appear together in a tweet they can be studied *relationally*, which is the approach proposed by Marres and Weltevrede (2012) who study the 'liveliness' of particular issues in terms of shifting hashtag associations. In activist circles however, when hashtags associated with a campaign (#occupy for example) it may be reasonable to assume a hashtag as a primary channel for information spread. But it is hard to see how hashtags emerge or battles between competing hashtags, or of course users who do not use hashtags at all.[172]

---

[171] So on one hand a human could design a script to chime in to certain discussions (triggered by the appearance of tweets with certain keywords) and even carry on rudimentary dialogue in a human way. Conversely when human users use the Tweet Button, which appears at the bottom of many webpages, they are usually invited to automatically Tweet something of the format '@Website 'Title of the Page' [URL]' which is remarkably similar to what a forwarding service would do. Add to this the difficultly that many accounts intermix manual tweets with automatic ones.

[172] Also there is much interpretive flexibility in hashtag use. Hashtags do not necessarily suggest a commitment to a sustained discussion, or any conversation at all, as in this case:

> PresentedByNT 7:00:53 AM
> #Hinkley #nuclear plant set to get go-ahead [URL] #nonuke #power #energy #propaganda #environment

Key Words

The above modes of spreading refer to specific platform features, which are more or less easily available for analysis, but it is also important to consider less obviously traceable ways of content spreading. The object of a user's search or trending algorithms need not be hashtags but can also be keywords. Thus users can draw in potential readers by a shrewd selection of terms, which might reflect what they think people are searching for (Murthy, 2013). Also Reider, in relation to the 'refraction' of links, points out that clever or humorous tweets may be shared more than others (2012). So it is important to consider, the potential relationship between the discursive content of a tweet and they way in which it is taken up or not by others. Even small modifications can contest or reframe a bit of information *and* simultaneously perform a material shift in the potential audience for that tweet. This is an important insight so long as it is understood outside of directly visible and traceable interactions and as a simultaneously discursive and material act – content and infrastructure are inseparable, they emerge together. What is discursively uttered transforms the socio-technical arrangements of the device and the drive to make content trend conditions possible discursive interventions. In other words it is almost impossible to spread content without in some way modifying it: *no amplification without modulation*.

Links

Each of these practices through which users come to view content (and either imitate it, retweet it or spin it) can potentially impact the flow of information. Yet many existing approaches, for practical reasons, bracket the question of which infrastructure is most important. These different content sharing practices must be considered together, but most means of data reduction give us only some of these practices at a time. A different solution would be to focus not on these infrastructures but on a particular object which travels.

Lerman and Ghosh (2010) also study 'information contagion' though networks but in a more interesting way. They instead study follower networks – which users 'follow'

This user is not primarily aiming at a response; they are employing hashtags to ensnare as many potential readers as possible. This use of hashtags relates more to metatagging: labelling content for prospective readers, rather that creating a space for discussion.

each other's tweets.[173] But rather than assume, these authors seek to *evaluate* the influence of follower networks on information diffusion, and they do so by isolating shared URLs in their corpus. One of the key features of Twitter is the ability to post content including images, video and most importantly hyperlinks, which are truncated using URL-shortening services. I will refer to this practice of posting links originating from mainstream media, blogs or alternative sources 'sharing' to note the parallels with similar practices other platforms like Facebook, though this takes on different forms in Twitter. By following a link, a stable object, they are in a position to judge the influences of network structure on sharing (the number of shares originating from a user's followers).[174] They find that around 50% of shares of a link result from follower connections, which begs the question – where does the other half come from?

In following Lerman and Ghosh's suggestion of scraping Twitter by links, it is possible to evaluate the centrality of these different sharing practices on a given links. Some of these practices are easily measurable like @ mentions and hashtags, while the influence of bots and keywords can only be inferred. But by reading the individual tweets and highlighting which content, hashtags, commentary etc. are taken up by other users, one can start to get a sense of what has the most impact. The study of links also follows the argument in Chapter II and III that it is crucial to understand what happens at the interface between platforms, how hyperlinks are understood differently.

Now grouping Twitter by URLS shared is just as provisional as other ways of circumscribing the data: the corpus will not of course include people replying to the share of a URL but not reposting the URL itself. Yet, by focusing on the link as an object helps remind the researcher that they are looking at a slice rather than a deceptively complete data set.

03. VISUALISING AMPLIFICATION AND MODULATION

[173] Follower networks have their own sets of shortcomings because they are constantly changing so it matters very much when the network is scraped.
[174] Lerman and Ghosh, however, oddly measure networks by followers of the first person to share the link only, as opposed to followers of followers. This assumption may have been more valid in the early days of twitter when most content originated from single users as opposed to a whole host of users and bots simultaneously.

So scraping based on particular hyperlinks provides a stable object through which to view the diversity of modes of information diffusion. With the help of DMI researchers I was provided with a spreadsheet containing every tweet featuring a particular URL[175] But how could this rather dense and repetitive list of tweets be analysed?

One obvious point of comparison to link-sharing, which comes from the study of controversies in science, is the deployment of citations in scientific papers. This is how scientists invoke outside facts, institutions, laboratories and devices which are difficult to challenge. A citation can be used as a jumping off point to build a further set of claims or the claims of the paper itself can be opened up and challenged. Latour (1987) refers to the qualification of citations as 'modalities'. positive modalities lead a citation away from its conditions of production (particular scientists in particular laboratories) and negative modalities draw attention to them. This would be an example of a positive modality which uses the claims contained in an article to make a further claim:

> Doing the math they have 8 days til cooling water boils off. /MT @AJEnglish: Power outage at Fukushima nuclear plant [URL]
> Tollie

It is however relatively rare on Twitter to challenge the article, to open up the black box of journalistic process to scrutiny: attributing claims to a PR source or question the terminology. This would be one example of a negative modality:

> RT @CarbonCounter_: Here's a man having an internal fight between his own dogma and the facts. [URL]

This presents the article not as an authoritative statement, but as the work of an author whose personal baggage is driving the content. But these kind of strong modifications, while important in terms of the articulation of controversies, are the exception in terms of link sharing. The more common tactics are very small modifications of basic tweet formats as below.

[175] However, this itself is a difficult task. Hyperlinks in Twitter are automatically truncated to accommodate the 140 character limit to a t.co link. Users also often truncate their own links with third party services such as bit.ly or tiny.url which means that a link in the data set could be a truncation of a truncation. Anne Helmond discusses this phenomena in a blog post: http://www.annehelmond.nl/2012/02/14/the-social-life-of-a-t-co-url-visualized/ (Accessed 28 September 2015). The developers of the DMI TCAT have design a script which resolved all the links to their original full version, which was used in this analysis.

Fukushima - Fear Is Still the Killer [URL]
Fukushima **--** Fear Is Still the Killer **- Forbes** [URL]
@nickbruechle Fukushima -- Fear Is Still the Killer **- Forbes** [URL]
#Fukushima -- Fear Is Still the Killer - **James Conca at Forbes** [URL] **#nuclear**
emphasis added to show modifications

Users (and bots) may offer an extra bit of punctuation, a hashtag, a mention or a brief comment but often the basic information is repeated over and over again. Yet even the lone hashtag, slogan or @ mention reframes the way the article is meant to be read *and* redistributes the link to a new potential readership: over time, certain phrasings (in the form of retweets or otherwise) become more popular and then fall away. Even these slight modifications may potentially alter the trajectory of a link's meaning and its diffusion depending of course if these modifications are taken up by other users, amplified or modified further.

I will use 'modalities' more generally to refer to all modifications of a tweet, including the deployment of '@'s and '#'s as well as discursive contributions.

Socio-Technical Graphs

Tracking these slight variations in the content which re-frame or re-distribute the link is a serious challenge for the qualitative researcher combing through individual tweets and the above approaches (participation framework, modalities) based in micro-sociological approaches give little indication how these practices may affect the macro dynamics of links over time.

What might help is a basic tool, proposed by Latour and Tiel (1992) for charting the dynamics of scientific controversies, which they called 'socio-technical graphs'. Given a collection of different accounts of a controversy, the researcher could code the different actors introduced into the account as arbitrary letters:

A New Nuclear Power Plant at Hinkley will create 25,000 jobs
A New Nuclear Power Plant at Hinkley will create 25,000 jobs for the French
Nuclear Power Plants produce Nuclear Waste

Nuclear Plant, Hinkley, Jobs

Nulcear Plant, Hinkley, Jobs, France
Nuclear Plant, Waste

ABC
ABCD
AE

When viewed in aggregate as the simplified letters, this tool can be used to identify which actors are most contested or deployed the most frequently and, if arranged chronologically, chart the rise and fall of different actors in accounts over time. As discussed in Chapter III, this is a way of reading texts in terms of the actor-worlds they present: how they articulate a controversy and what they leave out and how this might change over time. It also however in the sense of hashtags or @ mentions also draws attention to the infrastructure of the platform itself, the mechanisms of diffusion. Although Latour and Tiel's tool is intended for different kinds of texts (interviews) and manual coding, the basic premise could be applied to Twitter.

There are several ways of implementing this as an automated tool depending on the research aims. Rather than assigning a letter, the tool could colour code individual actors, both human and non human. This could be approximated automatically by harvesting proper names using a service like Open Calais or Alchemy (Venturini and Guido, 2012), though this may presume on the researcher's behalf what counts as an 'actor': humans, institutions, ideas? Also Open Calais or Alchemy would resolve different names, spellings of an actor into one entity, when these divergences of these articulations may be consequential for different articulations.

A less presumptuous method would be to merely highlight *unique* content (proper names or otherwise) and colour code it according to the user that first used it in a particular stream of link shares (see Image 11). This follows closely the manual work I first attempted to make sense of the repetitive streams of information and tracing the rise and fall of particular modifications. This helped me identify which account was responsible for content which gets appropriated later by others. It also helps identify even minor differences between tweets.

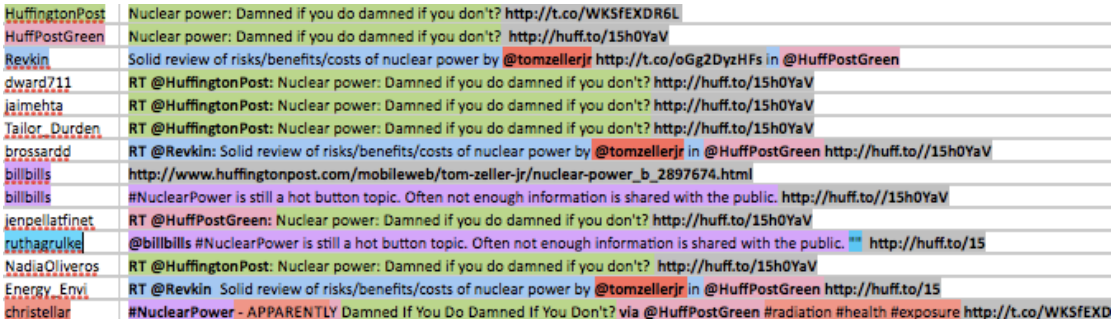


Image 11. Manual colour coding of Tweets based on user who first introduced the content.

When zoomed out, one could easily see which retweets are most prevalent and also which aspects (hashtags or phrases etc) retweets are composed of.[176] The downside is that the tool privileges originality and the order in which tweets appear which is helpful for retweets but completely meaningless in the case of RSS bots which operate independently of each others actions – thus the order in which content emerges is less relevant.

However, to proceed with the investigation, I started with a much blunter method. In the below image (Image 12) an Excel formula was created which searched through the sequence of Tweets, assigning numbers to *unique* Tweets, starting with 1 and ascending. If a Tweet was repeated, as in the case of exact retweets, the formula would assign the number of the first instance. Each number was then given a unique colour based on a subtle gradient (in this case from dark orange to yellow to green). If it is an imitation, the colour remains the same as the original, if there is any innovation, the colour advances by one shade.

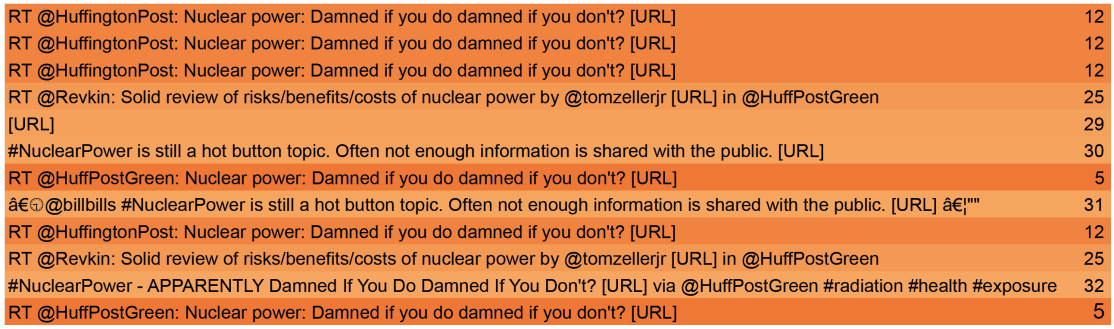


Image 12. Automatic coding of identical Tweets (defined by number on the right) colours gets lighter as distinctly new modulations are added.

[176] Johannes Passman's group at the 2012 DMI summer school attempted something similar with a view to identifying patterns of behavior as bot-like or human-like.

This simpler method does, as I will show shortly, give enough of an indication, in the sense of quali-quantitative methods, of the relationship between individual tweets and patterns of tweeting over time and, through comparison, the relative impacts of these discursive material behaviours on the success of particular links in quantitative terms.

04. THREE LINKS

In this section I am going to use this tool to analyse some URLs related to the particular controversies discussed earlier. The point of this analysis is to get behind the frequencies and numbers to the mechanics of how content spreads and the key insight that content changes as it spreads and particular sorts of content, due to the interaction of conventions with technical possibilities, have more opportunity to spread further.

Using the DMI TCAT interface (Borra and Rieder 2014) I obtained lists of every URL shared more than once for each of the days 18 – 20 of March, when the controversies in question were happening. The radioactive fish was discovered on the 18th, the same time as the blackout was happening, planning permission was approved for Hinkley Point C on the 19th and on the 20th it was discovered that the culprit for the blackout was a dead rat. This was limited to Tweets containing the key word queries mentioned earlier but I also checked to see if certain shares of the URL on Twitter had not been captured by these queries.[177] On these days I analysed all URLs in the data set directly dealing with nuclear power that were in English, about 20, and sampled the remainder of the links to confirm that they did not deal with the issue of nuclear power or the events of this week.[178] [179]

[177] I checked this using http://www.sharedcount.com/ to see how many times the URL was shared to see how many had not been captured by my key word query.

[178] Other URLs in the dataset discussed nuclear weapons programs in Iran and Korea; a soldier giving up nuclear secrets to his Chinese mistress; a Google Street View prank with a road sign pointing to 'secret nuclear bunker'; a nuclear alarm system error in a Chicago Metro and countless videos of teenagers playing the first person shooter game Black Ops which has a mode called 'nuclear'. This shows just how much noise there is on Twitter in any key-word based data set. But it is important to take seriously this noise because even a story which appears to be about Iran's nuclear program may be made to relate to its claims to be making domestic nuclear power and thus the controversy at hand. It is in fact very important for anti-nuclear activists to constantly reinforce the original link between nuclear power plants and bomb making - nuclear plants were originally only for producing weapons grade plutonium from uranium - generating power was an afterthought.

[179] Language will always represent an arbitrary boundary on social media research as many users interact across national and language boundaries all the time. Some very interesting links

I analysed each of the articles and, as far as possible, traced the article back to press releases and original sources to get a sense of how these, mostly online news articles, were first mediating the controversy before then considering how Twitter users were further modulating and mediating these articles. This was important to understand what was at stake in Twitter's distinct contribution over and above the news in articulating / publicising the particular controversies in question: either the Hinkley announcement or the various events at Fukushima. I also read comments underneath the articles which sometimes tied into the Twitter discussions.[180]

Of the URLs analysed I will concentrate on three, which I think contain key features of three modes of sharing I will describe: grassroots, broadcast and spin. In each case I will be looking for how users deploy @, #s, RSS services, and discursive modulation of links, as well as how these users self-present themselves on user pages and what kind of modes of diffusion likely result from this behaviour.[181]

---

were shared in Japan, especially originating from the TEPCO website. But a different dataset with Japanese words would be needed to study these stories.

[180] I also analysed some editorials with a broader focus including *Forbes* – 'Fear is the Killer' and Huffington Post's 'Damned if we Do Damned if We Don't' but I will not discuss these in my detailed analysis because I would rather focus on the events in question rather than this very different kind of commentary – which it should be said, sparked surprisingly little response on Twitter.

[181] It would be difficult to confirm the accuracy of the way these accounts present themselves, but for now I am merely interested in the relationship between self-presentation and tweeting behaviour.

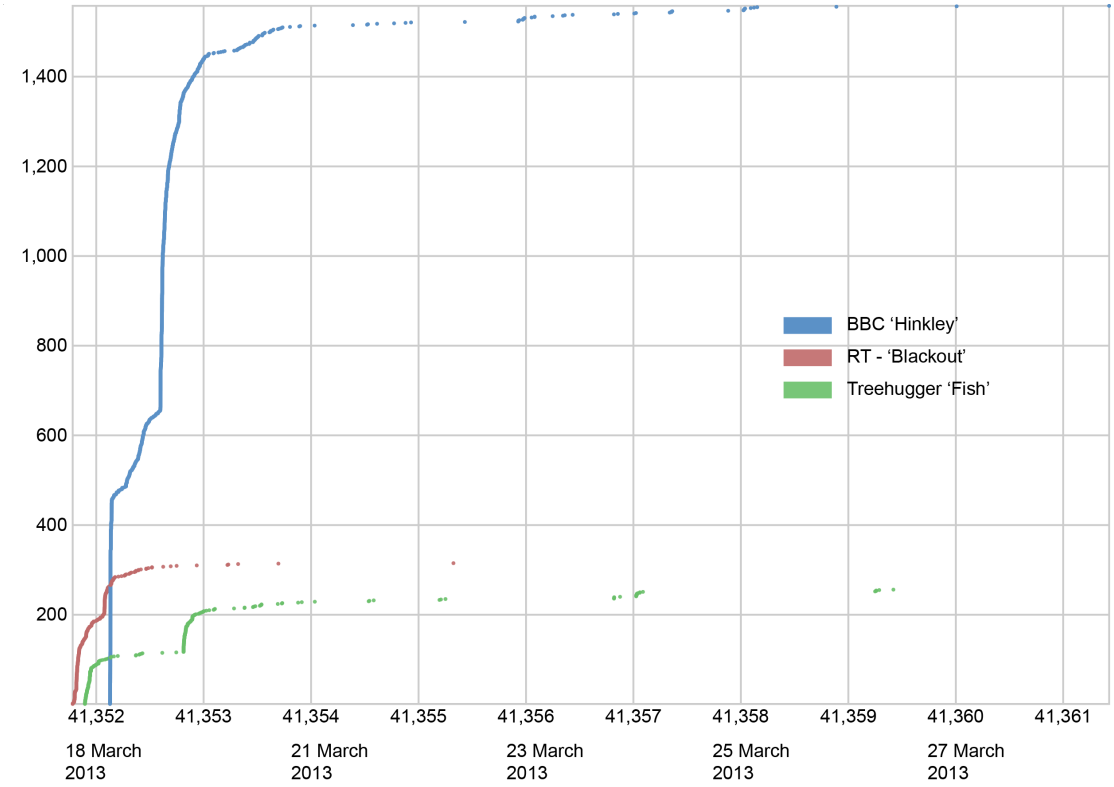


Figure 14. URL Frequencies https://goo.gl/ZVqm7O: Scatterplot of shares of three URLs over time, y-axis is number of tweets, x-axis is time. The Treehugger Article was shared 256 times starting on the 18th March and continuing for several days after. Russia Today was shared 310 times, starting on the 18th of March, peaking later that day and dropping off sharply on the 19th. BBC was shared 1,500 times, peaking on the 19th but continuing to be shared for more than a week.

Frequencies over time

I first found it helpful to create a frequency graph of each of the four URLs. They all have similar trajectories, rising quickly within a few hours and petering out with a long tail. The BBC story, which 'trended' the most seems to have two subtle bursts of activity, when the line becomes nearly vertical, while the Treehugger article has a more pronounced slump and reactivation about a day after first being shared. But as I will show these two bursts of sharing activity, that is in terms of 'liveness', have very different explanations, something which comes into more focus with the graphs, which also visualise shifts in content or 'liveliness'.

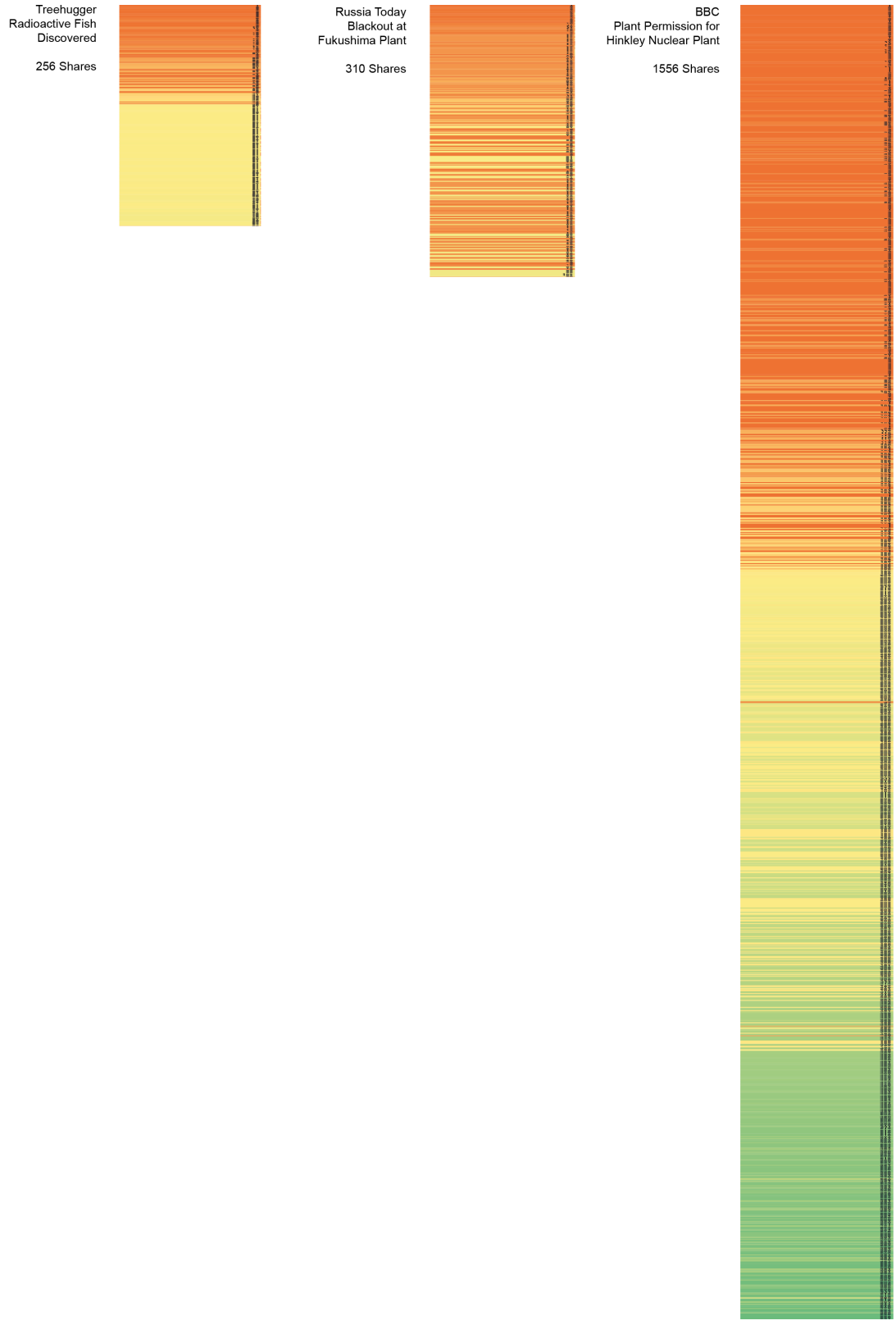


Figure 15. URL Ribbon Graphs https://goo.gl/uqmJcj: Each horizontal line is a tweet and the vertical length of the ribbons represents the total number of shares. Colours advance along a spectrum from orange to yellow to green. Progress along this spectrum indicates more modulated material while continued use of orange indicates recycled tweets from the beginning.

Looking at these sequences of tweets from a zoomed out perspective, one can get a rough sense of the dynamics of link modulation. The left-most ribbon (from *Treehugger*) makes a fairly uniform gradation from orange to yellow which suggests that most of the Tweets were modified or heavily personalised (by humans or bots). But there is an abrupt shift in the colour, and thus the content, where the orange coded tweets drop off completely – this is the burst visible in the frequencies graph. In contrast the Russia Today story is highly variegated — the orange tweets recur throughout – suggesting that the same tweet formats are being used over and over again. This may suggest the presence of forwarding services which offer limited options for customising.

The BBC story is particularly interesting because, like the Russia Today link, it is relatively uniform in orange colouring at the beginning, which means that the same tweets are being recycled (many being the first few tweets) but then it makes a completely abrupt switch where the orange stops. This is the second burst on the frequencies. In the final third of the BBC strip, the colour forms a more even gradation from green to dark green, again suggesting more modification. So these two shifts in behaviour represent both a shift in *content* as well as an increase in the rate of tweeting. The question then becomes: are these shifts in content due to revelations and developments in the controversy itself, are they because of dynamics in mainstream media, or are they coming from Twitter itself – either in the form of active reframing or more passive dissemination of the story.

In the following narrative, all times are in GMT and follower counts, which could not be scraped, are estimated based on follower counts at the time of analysis in late 2013.[182]

Treehugger – Radioactive Fish

Several sites picked up on an announcement made by TEPCO, the energy company in charge of the Fukushima plant, that a fish had been captured in their nets with

[182] URLs in the Tweets were converted into the text '[URL]' so that the graphs could treat alternative truncations as the same. A column from the DMI TCAT data set called 'source' allowed me to view which device the Tweet originated from. Twitterfeed and delivr.it for example are known bots but there are too many others to count.

unusually high levels of radioactive Cesium 137.[183] One version of this story appeared on the website Treehugger, an independent lifestyle magazine for environmentalists. The article, referencing a story in the English language *Japan Times*, reports that TEPCO has been catching fish near the accident site to monitor radiation levels in the ecosystem (the previous record was 500,000 Bq/kg but the new fish contained 700,000 Bq/kg. The Japan times story contextualises the number as 7,400 times the government limit for human consumption – but of course, these limits are themselves a controversy.[184] But while the author accepts the arrangement of facts, actors and institutions related to the fish he uses these to emphasise Fukushima as an on-going, rather than past crisis.

> 'A becquerel is a very small unit of measurement, so even 740k isn't as high as it might seem (though still very worrisome), but still, this shows that theFukushima saga is far from over even 2 years later...'

So how does this article spread on Twitter? The first Tweet is delivered at 9pm London Time (4pm New York time) by the article's author Michael Graham, and then by the website account itself. Graham has 9,000 followers whereas the website has 250,000 at the time of writing. The Treehugger Tweet sparks a minor rally of about 30 retweets, most delivered manually or with the aid of a twitter client, rather than bots. Whether they are retweeting the original tweet or using RSS, the message is nearly identical: article title, lead-in (truncated) and URL. The appearance of the lead-in 'Even two years later we are frequen..' helps emphasise the key message of the author. It is thus very important for journalists to consider both the title and the first few words of the RSS description in terms of social media coverage.

An hour and a half later, the UK branch of Treehugger then chimes in by simultaneously retweeting 11 of the users who already shared the story in this format:

> RT @EcoPassport: Fish caught near Fukushima contains record levels of radioactive cesium http://t.co/B0OhWBbY0S http://goo.gl/8kJBI

[183] Cesium is significant because it indicates *internal* absorption of radiation through water/food in the ecosystem rather than *external* absorption through exposure. Internal absorption has, at least historically, been downplayed by the nuclear industry and their scientific supporters because it indicates long term consequences for nuclear accidents like Chernobyl.

[184] To clarify this point, the author interestingly refers the reader to a Wikipedia article about Becquerel's, which is of course not without controversy in it's definition. See also Nowotny and Hirsch (1980) for a similar controversy.

Treehuggeruk 18/03/2013 22:37

This gesture, both acts as a 'thank you' to readers for retweeting and also, strategically informs the 11 accounts (and potentially their followers) that Treehugger has a UK branch which can be followed at this handle. It is perhaps no accident that all of these particular users originally failed to acknowledge the '@Treehugger' handle. This strategic use of @ mentions are ways of growing and maintaining networks of potential followers: networks are as much an effect of information spread as they are a cause.

But the story really takes off at 7:30 PM the following day when @GreenPeace retweets the story, GreenPeace has 800,000 followers at the time of writing and this illicits nearly 100 retweets in the next few hours and another 30 retweets over the following week. Nearly all of these appear to be manual retweets: meaning a user saw the Tweet and cut and paste it into their status.

The narrative of this particular link seems to support a very networked understanding of information spread, the sort that I was sceptical of earlier in this chapter. Most of the Tweets are essentially the same (few with commentary or hashtags) so they are not modulating the original message much at all. The amount of shares appears to be less about the discursive content of tweet and more about the number of followers a particular user has. GreenPeace has double the followers of Treehugger and produces more than double the results. It is their structural position in follower networks that seems to be most important.[185] The use of @ mentions is mostly attributional rather than a form of back-and-forth conversation or soliciting.

This is proto-typical of what I will call the 'grassroots' mode of sharing which seems to be most prevalent within loose associations of activist users (with the occasional large organisation like Greenpeace intervening): information travels in a networked fashion and is not substantially modulated, though hashtags like #nonuke and #green are occasionally deployed to alert particularly environmental *audiences* – most of the users who share the link or retweet these accounts explicitly mention environmental causes in their user pages.

[185] This seems to confirm Lerhman and Ghosh's conclusions about network influence but also shows that the first Tweeter is definitely not the most important.

So through a consideration of the interaction between conventions and strategies related to acknowledgement, and technologies like the @ mention certain sorts of contents tend to spread in particular ways, which has consequences for how much a link gets shared cumulatively.

Russia Today – Blackout at Fukushima

The second article I will discuss was also directly triggered by a TEPCO announcement, though this time, a more pressing one regarding another loss of power to the reactor. This was picked up by a number of sites concerned with the energy industry, but less so by the mainstream news. One of the most shared articles on this topic came from alternative cable news network RT (Russia Today). The article itself is a fairly straight reproduction of TEPCO's press release, detailing the reactors affected and the claim that the plant can go for four days without power before the water temperature rises to dangerous levels. But the article does make the significant step of including a link to another Russia Today article about the radioactive fish. This is, most likely, a means of directing traffic but also recasts the incident as one of a sequence of problems which again presents Fukushima as an on-going concern.

RT is a Moscow-based English language satellite television channel aimed at a Western market and an increasingly visible player in the social media sphere, especially, according to the profiles of the users sharing the link, with accounts that identify themselves as politically conservative. Yet a large proportion of the Tweets come from accounts, which present themselves, not as individual human users, but as alternative news outlets. For example:

> @ConspiracyR - Conspiracy Realism
> 24 Hour News that Informs you of world issues, Follow ConspiracyRealism for the Latest News and Updates and Subscribe to the URL below #NWO #HAARP[186]
>
> @UnreportedNews1
> Unreported News That Doesn't Make Headlines But Should.

[186] These hashtags are typical of 'truther accounts': #HAARP is High Frequency Active Auroral Research Program - communication program sending radio signals over long distances - conspiracy theorists blame this for many unexplained events. #NWO is the New World Order which refers to a conspiracy to form an oppressive world government

These types of accounts, which are either partially or entirely bot-driven, frequently express some scepticism toward the mainstream media and what they cover.[187] Some could be classed as 'conspiracy theorists' or to use their own positive self-identification 'truthers' and in the particular case of Fukushima, 'preppers' – people who are readying themselves for the end of the world, in this case as a result of the ongoing radiation leak from Fukushima.

These bots automatically share articles published in certain outlets on specific topics, like a specially curated magazine for their target audience. They are gatewatching (Bruns, 2005) the news but potentially doing so automatically through bots. As with the last article, the first few shares come from RSS bots including RT itself @RT_COM, which at the time of writing has 544,894 Followers. However the hundreds of bots which retweet a near identical message need not be followers of RT, they only need to be plugged in to RT's RSS feed. Thus, it is not uncommon for other Twitter bots to share the story even before the source account.[188]

> @Muschelschloss -TEPCO reports power failure at **#Fukushima** stops cooling system — 18.03.2013 - RT News http://t.co/BEX3ODlrGW cc: **@haloefekti**
>
> **@gabyverdier** http://rt.com/news/fukushima-power-failure-cooling-445/#.UUdl4beRZqQ.twitter
> *Emphasis mine.*

While some of these users who share links retweet each other in an activist mode, many rely on forwarding services, which explains why there was so much repetition in the earlier graph of this URL, in contrast with the treehugger piece. Because they are imitating news outlets, there is a lack of commentary on the substantive issues, they are impassively imparting the basic information. This mode of sharing, heavily reliant on bots or bot like behaviour, I will call 'broadcast'.

---

[187] As Latour notes, what is interesting about conspiracy theories is that they ironically accept the premise of the construction of knowledge (2004).
[188] Like many homespun news services on Twitter, @Muschelschloss prides himself/herself on not only covering the right stories but doing so very quickly: 'Sometime I am faster than #Reuters - Tweets in German & English'

Again the most frequent modifications merely reproduce the article title and lead in with minor modifications, mainly hashtags, to direct it toward a targeted readership — those who are sceptical of mainstream media coverage. But hashtags simultaneously can widen the 'participation framework', the potential audience, and at the same time re-articulate the content of the article.

> CitizenoftheWo4
> **#TEPCO** reports power failure at **#Fukushima** stops cooling system — http://t.co/8a5A5UHp8T **#Nuclear #Energy #Corrupt #GE #Reactor #Design** http://rt.com/news/fukushima-power-failure-cooling-445/
> *Emphasis mine.*

Tagging the proper names in the article title is a common tactic for soliciting readership. While the trail of hashtags on the end: '#Nuclear' and '#Energy' would direct the tweet towards users interested in these issues (potentially on both sides) '#Corrupt #GE #Reactor #Design' in addition function as charged commentary on the subject – General Electric was the company responsible for the, some say, shoddy design of the Fukushima reactor. This is a minor example but there are some tweets like the following, which more dramatically shift the content to include the media itself:

> @Harleypyrate02
> BREAKING NEWS>TEPCO reports power failure at Fukushima stops cooling system! http://t.co/0tVcIMHDZj >>>> **Any U.S Media On This???** http://rt.com/news/fukushima-power-failure-cooling-445/
> *Emphasis mine.*

This user is using the RT link to make the common observation that mainstream media is suspiciously silent on nuclear dangers. Again, this study of a link shows that certain sorts of contents are associated with certain modes of spreading: in this case articles which support a 'truther' perspective on Fukushima. But while the use of bots generally discourages modifications, the very act of naming and targeting particular audiences changes the presentation of the article – adding insights and connections which were not made explicit in the article. In the narrative of the next URL, it becomes more obvious what is at stake in the use of bots.

BBC – Announcement of Hinkley C

By far the most shared link during this series of days was a BBC story about the decision by the UK government to grant planning permission to French Energy supplier EDF. What is crucial about this particular article is the way it changes over time and how it is enmeshed with social media through RSS and other technologies. The BBC article is also significant because it contains a wide variety of actors gathering around it with different sharing practices, but the type of sharing I want to highlight – because it is most prominent in this article – is 'spin'.

The article, originally titled: 'Hinkley nuclear plant set to get go-ahead' as of 18 March 2013 (Last updated at 22:01 ET), cites a number of facts which are identical to those put forward in the eventual government press release: will deliver power to 5 million homes; 20-25,000 jobs during construction; 20 years since the last nuclear power plant.[189] All of this key information links the project to an economic articulation of nuclear power which de-politicises the issue, while the information (buried at the bottom) attributed to the Stop Hinkley anti-nuclear group, critiques the proposal based on their preferred terms of health and safety.

As it stands, this article, in supplying these particular facts and presenting these spokespeople, the actor-world of the article, favours an articulation of the Hinkley issue which is favourable to EDF and the government by presenting it in economic terms; but would this be challenged on Twitter? The first two tweets come at 3:05 am on the 19th from what appears to be a BBC bot. These tweets and the next 500 that follow, are nearly all from various forwarding services, mainly Twitterfeed, dlvr.it and sharedby. But while accounts sharing the RT article used relatively few hashtags, for the BBC article, this is the main way that links are modulated – either with the aide of a forwarding service or not. I suspect this is because the BBC article is mainstream enough that the presumed readership of the piece – quite specific for Treehugger or Russia Today – must be re-specified through Twitter.

One type of hashtag, whose significance may not be readily apparent, is the use of geographic hashtags such as #UK #Somerset #Hinkley. It matters considerably in terms of media framing the scope or scale of the controversy. If the planning decision is

[189] Although the article has no byline, this information is credited to John Moylan who earlier wrote an article broadly sympathetic to nuclear: 'Hinkley Point Nuclear new build could boost economy' a few days earlier (Available at: http://www.bbc.co.uk/news/business-21788883 Accessed 3 July 2015)

'world news' than it could become easily linked to contrasting energy policies in Germany and Japan. A national (#UK) issue however more easily relates the story to the energy demands of the UK generally and the economy, which might be preferred by the government. However, if it is a local issue and directed at residents of Somerset and Hinkley then it may be more easily linked to local concerns over public health but also directed away from the national level at which energy-based decision making happens (Johnstone, 2014). I am not in a position to speculate whether or not accounts are consciously channelling the story to certain readers in this way, only that these seemingly innocuous hashtags potentially impact both the content and potential readership of the article.[190]

When the actual announcement happens, sometime around 2PM, the title and content of the article are significantly updated and expanded to reflect a wider context. Thus the RSS-led stream of Tweets changes:

> BBC News - Hinkley nuclear plant set to get go-ahead [URL]
> 2:08:50 PM
> ----
> BBC News - New nuclear power plant at Hinkley Point C is approved [URL]
> 2:16:53 PM

The article title changes between the above two tweets.[191] This launches another deluge of RSS feed driven tweets featuring the new title: 'New nuclear power plant at Hinkley Point C is approved' (19 March 2013 - Last updated at 10:20 ET).[192] This explains the sudden shift in content in the ribbon graph earlier and reveals just how

---

[190] Also, it is likely that these feeds were set up to follow the BBC's own news categorisation via different RSS streams (e.g. – world, national etc). It is also possible that bots can determine the type of hashtag based on the appearance of certain keywords. So this type of framing may still be somewhat set by the BBC. Along these lines, one of the more common and striking hashtags for this story is #business because it situates the article within the economic frame preferred by the government and encourages users in industry or management to participate in spreading this article. But this is also an established category for news reporting:

> #news #business Hinkley nuclear plant awaits go-ahead [URL]

[191] It should be noted that versions of the original title were still being published by RSS services even days later.

[192] Fortunately in this case Twitter helps analyse the changing news article in a way which would be impossible otherwise. Unlike the Guardian, the BBC website does not allow access to past versions of an article, these are simply overwritten like a palimsest, though they often give a 'last updated' date to indicate that changes have occurred. but using the internet Archive one can see selected, though arbitrary slices of the article at different points. The earliest archived version has the following title, which anticipates an announcement the following day.

central technologies like RSS and services like Twitterfeed are to the dissemination of news information. Depending on how regularly the bots are set to check the BBC feed, by changing the title, the BBC may illicit two tweets from some bots. This is important because while each tweet changes the framing of the story, the story is also changing independently of Twitter, though potentially in dialogue with perceived social media chatter.

It is only after UK working hours that more proactive commentary begins to emerge in earnest. The proportion of RSS feeds declines dramatically and there is substantially more variation in the content, as one can see from the graph. In many ways it resembles a forum style discussion, like the article comments themselves, though with few actual exchanges between participants. I will refer to this mode as 'spin' after the political profession of 'spin-doctoring' or strategically manipulating media perceptions, though I do not invoke this term in a necessarily pejorative way.

According to the user profiles, this practice could be associated with users who identify themselves as individuals, as opposed to organisations, who generally do not identify themselves strongly with either environmentalism or the topic of nuclear energy, though there are some exceptions. Of this commentary there are a few different strategies which emerge. As with hashtags, users can substantively broaden the scope of the event, relating it not only to events in the UK, but to world-wide opinions on Nuclear, or by the same token de-publicize it:

> The Germans the Japanese and others have decided - no more nuclear... So what do this lot do? (Harrumph - the... [URL]
> FidoMorgan 2:31:10 PM

> BBC News - ''New nuclear plant at Hinkley Point C is approved [URL] What's the C stand for? Chernobyl?
> Bydgoszczanka 6:32:44 PM

The first user refers to the fact that the German government promised to phase out nuclear after the events at Fukushima and the Japanese people (but not the government) have become staunchly anti-nuclear. Occasionally as the day wears on there is a direct link made between Fukushima and Hinkley:

> After Fukushima, still don't understand why a nuclear power plant approved in the UK [URL] Fuck French polluter @edf
> Tweetstrike 5:07:02 PM

> Fukushima spent fuel ponds in danger of boiling dry and UK announces go ahead for Hinkley C [URL] Not ideal timing I think
> TonyJuniper 7:04:32 PM

These tweets are strategically making connections between disparate controversies around the world, in the sense described by Andrew Barry in his concept of political situations (2012). The first finds the possibility of new nuclear inconceivable after Fukushima while the later, which was retweeted 15 times that evening, emphasizes the on-going nature of the crisis. These tweets which spin the link often receive small but quick bursts of re-tweets, in some cases this may be due to the celebrity of the Tweeter – Dr. Helen Caldicott's minimal message is retweeted 32 times – or in other cases due to the perceived cleverness or substance of the commentary.

There are also plenty of tweets which celebrate the announcement of the plant but less frequently, and usually positioning themselves as adopting nuclear, perhaps reluctantly, as the *pragmatic* option.

> New nuclear power plant at Hinkley Point C is approved [URL] At least the British are being realistic for Energy sources.
> MaxwellMarshal 9:56:50 PM

> BBC News - Hinkley nuclear plant set to get go-ahead [URL] < 25k construction jobs clean reliable elec for 5million homes
> Kirstygogan 8:50:06 AM

@Kirstygogan's profile reads 'Climate, energy, politics, science. Communications director in UK low carbon electricity sector. Mama. Feminist. Views mine. London · uknuclear.wordpress.com' Although this user is tweeting in her capacity as a private citizen, with the common caveat 'views mine' the blog link reveals that she is a press officer for a nuclear lobby group, which she positions as a 'low carbon' energy source.[193]

[193] Just as with other technologies of elicitation (Lezaun and Soneryd, 2007), there is a certain extent to which Twitter performs the 'citizenness' of the users commenting.

With this URL, there is much more of an attempt to contest or modulate the substance of the article, rather than simply disseminating it, because it will be seen by a wide variety of actors, rather than a specialist audience. It is perhaps because of the trending potential of a BBC URL that users reflexively decide to engage on this register. But despite the creative use of hashtags and commentary, still the majority of the messages disseminate the article in a bot-like way (whether they are bots or not). This allows the BBC journalists to change the content of the story, including the title while maintaining a unique link. So even though this article was changed, it is counted as one article of 1500 shares, not two articles of 750 shares each in the metrics of any trending algorithms. Constant updating allows them to trigger RSS based bots to potentially tweet the story twice.

As I argued in the last chapter, while the quantitative accumulation of materials is a legitimate form of action, if we only evaluate the contributions of participatory media users in quantitative terms than we may miss the role of discursive commentary in not just making controversies visible, but making them visible in certain ways and to certain audiences.

Three Modes of Sharing

Twitter like all internet platforms, may direct and shape modes of participation and sharing, but it also affords a great deal of 'interpretive flexibility'. Van Dijck emphasises that Twitter was originally designed as a generic service (like a utility) which accommodates drastically different types of behaviour (2013). In the above examples I have identified three main modes of sharing articles summarised in the following table:

	Self-Presentation	**Device**	**@**	**#**	**discursive**	**diffusion**
Grassroots	Issue Activists	Web/phone	Attribution	Slogans	--	Networked
Broadcast	Truther / Alternative	RSS feeds	--	Tagging	--	Backchannel
Spin	Lone User / PR	Tweet Button	Soliciting	Framing	Comment	Small bursts

Table 1: Modes of Sharing

In the 'grassroots' mode, messages are disseminated in a networked way, almost like a classic letter chain, and are thus concerned with polite and careful attributions using @ with some use of campaign-oriented hashtags. Users employing a 'broadcast' mode rely on RSS bots primarily to automatically share articles, some of which include generalised hashtags like #news, #environment with the goal of widening the participation framework, or human users act similarly in a bot-like way. This results in massive eruptions of tweets timed with the release of content. Tweeters practicing 'spin' are often much slower but leveraging the article in the service of a message, using Twitter like news comment sections. The goal is to attract limited bursts of retweets either ratifying or less frequently, arguing with the comment.

Now although I have selected articles, which perhaps due to their readership generally are characterised by these modes, these divergent practices are always being mixed, even in the same Tweet. In the case of the BBC article, all of these modes intermingle in interesting ways, and it was only later on that the spin mode began to emerge. Though both the Treehugger and the BCC article had matching bursts of sharing activity, the former peaked because a highly networked actor (Greenpeace) tweeted, while the latter peaked because the author updated the title, which triggered the RSS bots.

The point in highlighting these divergent practices is that the game of making content visible on Twitter is again slanted to bigger and more technically savvy actors, but this only becomes clear when interactions between content and the mechanism through which content travels are considered together.

05. THE MODULATION SEQUENCER

The very crude approach I used above delivered some significant insights into the process of trending which would have remained more obscure had I aligned myself with the platform features easily offered up by the device. This approach originated from a specific set of controversies and a particular set of stakes: how could social media balance the so called 'mainstream media' account of the controversy. In the remainder of the chapter, if the reader will indulge me, I want to push these insights further by reworking the tool based on what I have learned. If qualitative analysis and quantitative tools are meant to be brought together, they should also feed off each other in an iterative process.

As a by-product of the analysis, it was discovered that that information does not diffuse outward from a single source, as in many of the networked or evolutionary metaphors – several completely independent trajectories emerged, some based on bots or retweets and more rarely conversations. So rather than rest on this metaphor of visible, traceable contagion, it is important to detect distinct typologies of tweets (retweets or particular bots) and then look for the modifications within them.

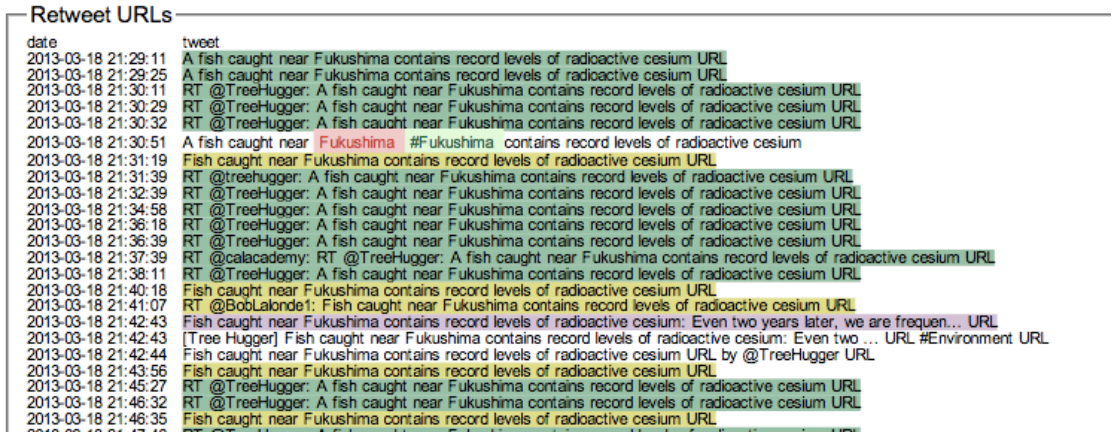
This could be detected automatically,[194] but for the purposes of this investigation it was simpler to identify typologies of Tweets around blocks of identical text. This involved ignoring some of the infrastructural elements such as the particular names of retweeted users 'RT@____', which would be constantly changing, and automatic truncations, '…', leaving only the basic text to identify typologies with. Each typology was then given a unique colour. In the below image (Figure 16), the different typologies are first highlighted in the leftmost column, which contains all of the tweets in time order. Then, for readability, each colour coded typology is given a separate column to the right starting in the order in which they first appear.[195]

Figure 16. URL Sequencer Treehugger https://goo.gl/kJelcX: http://www.treehugger.com/energy-disasters/fish-caught-near-fukushima-contains-record-levels-radioactive-cesium.html

[194] This detection could take a couple of forms. The Levenshtein distance is an algorithm which detects changes in words. Put simply, it measures the number of characters which need to be changed in order to turn one word into another. So turning C-A-T into M-A-T-T-E-R would be a distances of four: turn C into M and add T-E-R. The same logic could be applied for words in a short phrase like a Tweet and this could be used to cluster types of Tweets.

[195] Thanks go to Erik Borra from the Digital Methods Initiative for suggesting this mode of presentation.

If tweets are truly unique (most often instances of 'spin') they will appear in the leftmost column without highlighting to aide their identification. If one hovers over a particular Tweet, terms or hashtags that were added to the basic tweet text are highlighted green and words that are removed are highlighted red. In the below example, the original tweet 'A fish caught near Fukushima contains record levels of radioactive cesium' was modified by placing a # in front of Fukushima.



Detail of Figure 16 – Treehugger article

The zoomed out view gives some indication of the dynamics of when certain types of Tweets arrive and could be used to more easily profile the trajectories of particular types of links. We can see from the above graph, as discussed earlier, that Treehugger was dominated by the retweets of a Greenpeace link (in pink).[196]

In the next section, I will apply this tool to more quickly analyse a handful of links in the later part of the year, based around another series of media stories and announcements.

Strike Price Announcement

I showed earlier in the chapter how the announcement of a new nuclear power plant, overshadowed a series of further complications at Fukushima. Twitter users and alternative media outlets of course made the link between the two and also modulated the presentation of the link to benefit their cause. In this section I want to discuss two

[196] Currently each Tweet receives it's own row in order but these rows could also be spaced according to timestamp to get a better sense of the dynamics of sharing.

further controversial events. Again, these were identified by first monitoring the nuclear key words for spikes in the frequency of activity and then zooming in on particular moments to investigate the potential controversy. The first controversial event was another announcement driven by the government and online news outlets but the second represents a very different sort of intervention originating from Twitter itself.

On the 20th of October, Hinkley re-emerged in the news as a strike price was finally agreed between the government and EDF. The strike price was an agreed price for electricity generated by the plant which was set at £92.50 per mega watt hour (MWh). Large infrastructure projects are often built with government bonds or future tax revenue but this private endeavour was being financed by guaranteed future income. This was essentially a bet on the price of electricity – if the market rate goes below this – the government pays EDF the difference, if it goes above it, EDF reimburses the government. The announcement was another carefully orchestrated 'pseudo event' launched simultaneously through several media channels: energy Secretary Ed Davies gave a speech in the house of commons, a press release was placed on the government website[197] and a series of mainstream media stories rehashed the press release (just as with the announcement of planning permission in March). One particular BBC article was shared 656 times, despite the fact that, significantly, this news was categorised by the BBC as a lead article in the '*business*' section, while the planning permission article was national news. Despite the fact that this was a far more decisive development than planning approval it receives about half the Twitter traffic, perhaps because of this categorisation.

The article itself firmly sticks within the economic articulation of the controversy: it leads with the information that the plant will be built with a consortium of French and Chinese companies and, as the government emphasized, not by the British taxpayer. Although the story clearly makes the point that the strike price is currently double the cost of energy in the UK, it also includes a chart directly from the Department of Energy and Climate Change (DECC) which shows that this strike price, which takes effect in 2023 is lower than the *current* strike price for renewables which are all over 100 per MWH.

[197] Available from https://www.gov.uk/government/speeches/agreement-reached-on-new-nuclear-power-station-at-hinkley (Accessed 30 Aug 2015).

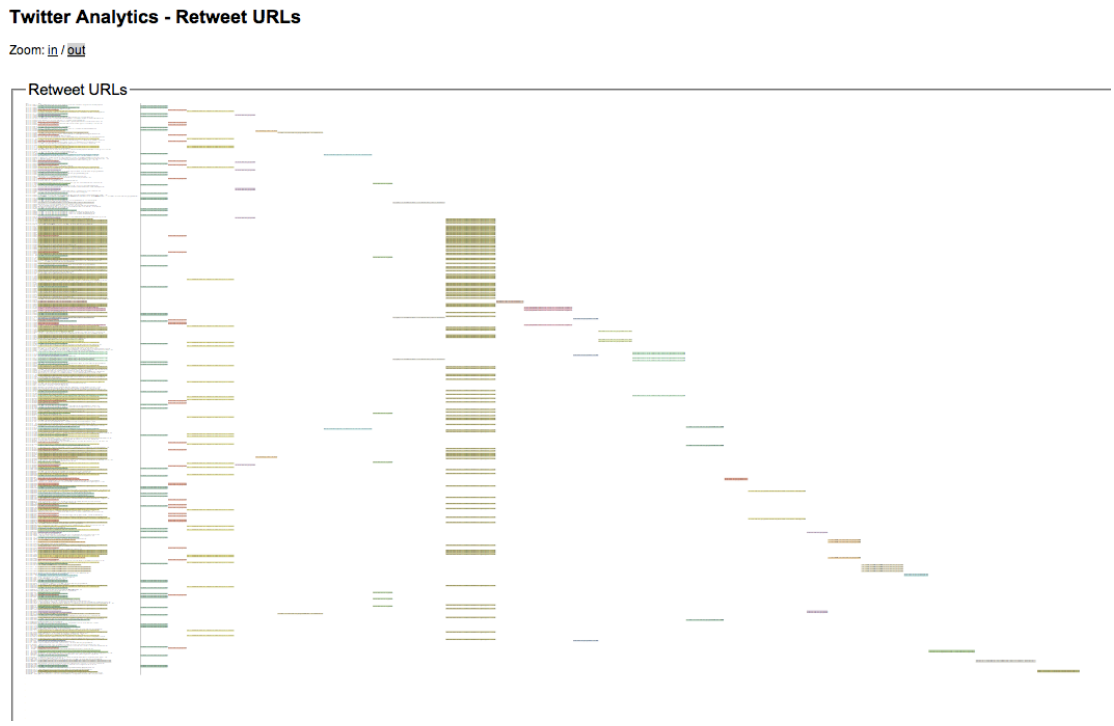


Figure 17. URL Sequencer BBC Strike Price article https://goo.gl/mPke8E: http://www.bbc.com/news/business-24604218

One of the largest retweets of this story (in greenish brown above) draws attention to the economic case against renewables:

> How nuclear power compares to renewables on wholesale cost of generating electricity £MWh

But as several Twitter users and commenters point out in the other Tweet typologies, it is misleading to compare a *projected* nuclear price to the cost of renewables now – because the cost of renewables is very rapidly decreasing with new technology and will likely be much lower than nuclear by 2023 (Dorfman et al., 2012). As with the BBC link discussed earlier, the dominant articulations of the article are based on those given in the article: Twitter users may either accept or reject these articulations, but their responses are, for the most part, necessarily in relation to them.

The 'Twitter Storm'

But it is not always the case that activity on Twitter is led by online news or blogs. Looking through the dataset to the spikes in activity, most are explained by either

revelations relating to Fukushima or relating to announcements concerning Hinkley and EDF, or indeed as is the danger with keywords: the crisis over North Korea's nuclear test or negotiations over Iran's nuclear programme. But one extreme spike on the 17th of November 2013 turns out to be a reflexively generated 'Twitter storm' or the deliberate engineering of 'trending' content.

On 23rd October, a link to a page on Pastebin accumulated a lot of shares: hundreds in a day and 3,252 shares over the coming weeks. Pastebin is a document hosting website normally used by programmers to store and share code but this time it is a different sort of script (Latour, 1992). The text talks about the deteriorating situation at Fukushima and the mainstream media's relative silence on it and proposes a 'Twitter storm' on the hashtag #Fukushima, giving the names of several accounts to follow and retweet. It also links to a live countdown to the 25 October at 12 CET 20:00 Japanese Standard time.[198] The original user @wattashit3, Tweets the link at several Anonymous affiliated Twitter accounts, including one, Lorax, which is a open account for anyone to use: the password to the account is given in the public profile: (https://twitter.com/Doemela). A leading hacktivist called @Amarandrill, retweets it, causing many of his followers to do the same, but the link is mainly driven through the army of Anonymous related accounts 'AnonNews' 'AnonPunkZ' etc.

Although there is a definite spike in activity for the hashtag Fukushima on the 25th, the organisers are not satisfied and change the text of the pastebin file to then refer to the 17th of Novermber – giving it more lead time to build momentum. This time the hashtag #Fukushima is overrun by an all out assault of grassroots, broadcast and spin.

Users start to lead it as the countdown approaches:

> RT @fukushima_actu: TODAY #FUKUSHIMA #TWITTERSTORM â˜¢ http://t.co/tlNYPZIr5V â˜¢ http://t.co/UxNXjXsR6p â˜¢ http://t.co/rFFr9IhMVt 10 25 4:57

One can neatly distinguish many of the Anonymous led activity in this hashtag because they tend to type #FUKUSHIMA in all caps, as is often the Anonymous style.

---

[198] Available at: timeanddate.com/countdown/generic?iso=20131117T20&p0=248&msg=%23FUKUSHIMA+TWITTERSTORM (Accessed 3 July 2015)

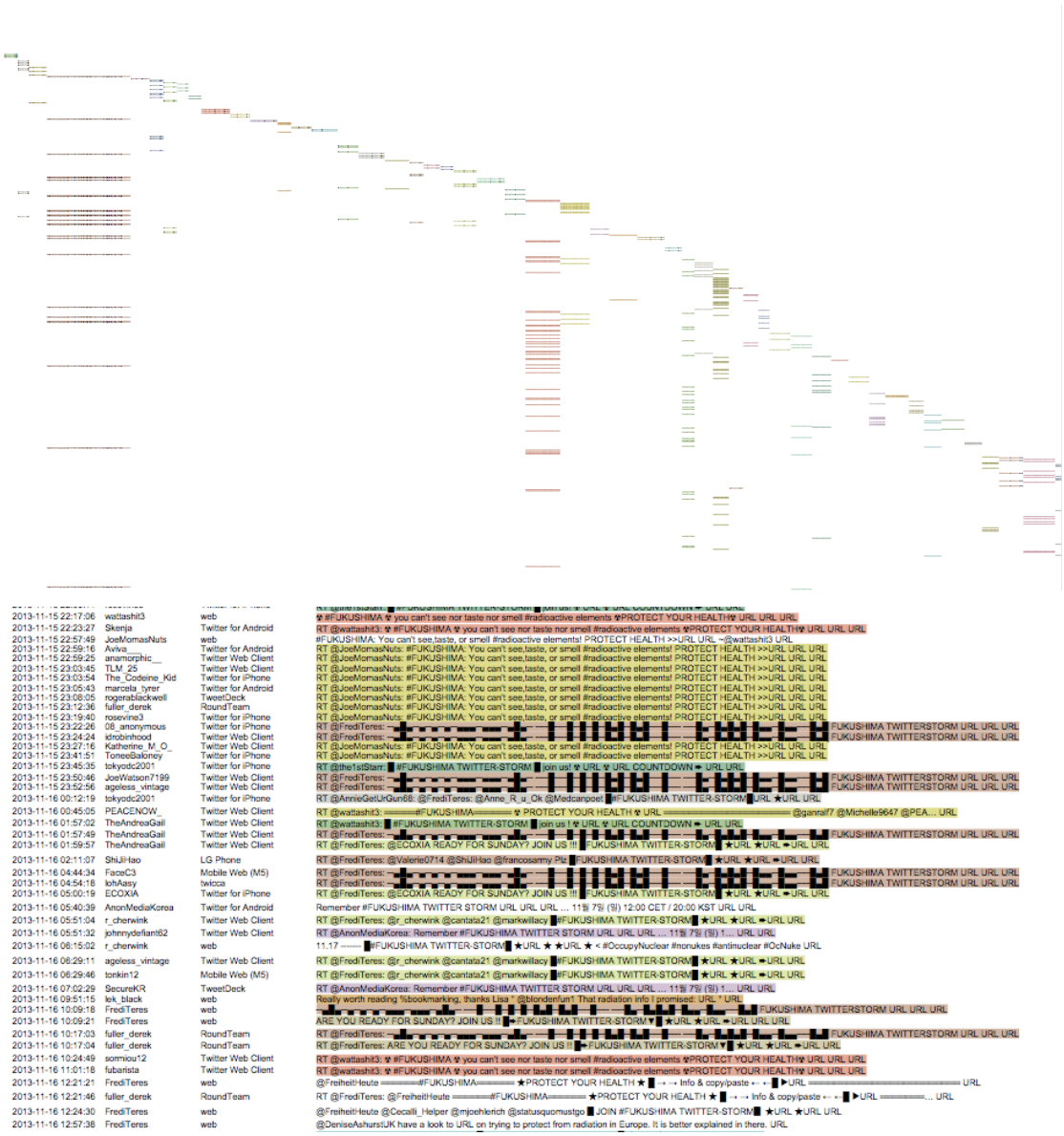


Figure 18. Sequencer 'Twitter Storm' https://goo.gl/05hV9G: http://pastebin.com/8qQTexQF

In the URL sequencer (Figure 18) we can see many variations on the link, mostly through short retweets through Anonymous's follower networks and the occasional bot. The highly diffuse graph shows a very ad hoc and unpredictable spread of information in contrast to links which spread through other more established networks.

Again, users are encouraged to retweet tweets from a selection of vetted users.

> RT @wattashit3: @FrediTeres #Fukushima #MSM black out #Gundersen 'Only thing saving us is internet' http://t.co/6Ft1UgmPGM @ENENews http:â€¦ [[http://enenews.com/columnist-fukushima-a-major-global-threat-to-all-living-flora-and-fauna-mainstream-media-is-n

Many users share the above article from Enenews, an independent energy news website), which is a rehash of a, now deleted, editorial in Toronto's Caledon Express. The author warns of the severity of the disaster for 'all living flora and fauna' and exclaims, because of the failures of the 'mainstream media', that: 'Only the internet can save us'[199] [200]

Ironically, however, there is little evidence I could find that the Twitter storm was actually trending in any national Twitter front page. Although a lot of activity was generated, it did not appear to leave the sphere of Anonymous, 'truther' circles. There were as far as I can tell, no mainstream media stories about the Twitter storm and ironically, the only reference to it is actually in a story about an *offline* event. This instance of 'digital demonstrations' occurred in tandem with a physical protest by Anonymous at a (old media) television station in California.[201]

> RT @ENENews: #Anonymous protests lack of #Fukushima coverage at TV stationâ€'Show Host: I wonder if media paid not to report on it http://t.câ€¦

As I have repeatedly shown in this chapter, Twitter is largely a responsive medium reacting to mainstream news stories by discursively reframing what is said and materially shifting the audience for a particular story at the same time. However Twitter can also generate its own content, but ironically it may be possible that something like a Twitter Storm only really begins to trend when other media such as online news recognise it and legitimate it.

---

[199] Available at: http://enenews.com/columnist-fukushima-a-major-global-threat-to-all-living-flora-and-fauna-mainstream-media-is-not-the-best-source-of-information-gundersen-the-only-thing-saving-us-is-the-internet-audio (Accessed 3 July 2015)

[200] The anonymous led Twitter storm does account for most of the activity on these days but it also coincides on the 17th with a few mainstream stories finally detailing the difficult clean up efforts and how this leads TEPCO into uncharted territory.

[201] Available at: http://www.kmph.com/story/23988783/anonymous-stages-mainstream-media-protest-at-kmph (Accessed 3 July 2015)

06. CONCLUSION

This chapter has described a series of controversial events in the UK, and also worldwide debates about nuclear power, through the sharing of URLS on Twitter, but the main point I wanted to make concerned the relation between content and the spread of information. In the case of Twitter at least, one cannot separate the content of media from the infrastructure or through which they spread. There is no spread of hyperlinks without modification and modifications affect the potential spread. This separation however gets enforced by methods which either consider the infrastructure or the content in isolation, or rather, presume that particular ways of spreading are fixed. This is an especially important point to make because one of the primary functions of Twitter in everyday sort of public science controversies is to disseminate information and it is important to consider that Twitter does not merely amplify or whip up controversies but can subtly alter their framing in the process.

This is, potentially, important because, even though the nuclear controversy in the UK played out, largely out of the public eye or public hearings though the negotiation of a strike price, it was crucial for the success of this economic device that it be consistently presented in favourable economic framings (in relation to job creation and growth, and specifically confined to the national context), something that Twitter users disputed consistently. Although there is little evidence that social media interventions had any significant impact in this regard, it was clear that both government and industry and actors like Anonymous view Twitter as an important space for influencing and conditioning more mainstream media representations.

As in the other chapters, I found that Twitter is not exactly a level playing field. In the first sequence of events, the mainstream news, which published a story quite favourable to the interests of the government and nuclear industry, dominated the social media landscape at the expense of potentially provocative stories about the continuing effects of nuclear accident: an economic framing of nuclear power succeeded over a health and safety one. This of course reproduces old power asymmetries between mainstream and alternative sources but it also has to do, at least in part, with follower networks, technical features of articles (use of RSS and Permalinks) and the strategic integration of comments. However, through the creative use of hashtags, @ mentions and discursive commentary, Twitter users were able to occasionally spin the meaning and significance of these articles and direct them

towards relevant potential audiences. So while the BBC story 'won' in terms of total shares, as I argued in the last chapter, we must also appreciate the more subtle, symbolic impact of re-framing and modulation.

This chapter built on the methodological gains of the previous chapters to propose an approach in which quantitative and qualitative techniques worked more in concert. Firstly, following the work of Chapter IV I (with the help of programmers at the DMI) was able to extract URLs which were not readily available from the API. In Chapter V, I tried to juxtapose numerical and textual data and, in this chapter, I was able to bring them closer together so that reading and viewing effects in aggregate was simply a matter of zooming in and out. The act of discursively analysing Tweets was also augmented though various types of colour coding revealing relationships between discreet tweets. Crucially this solution lives up to the promise of quali-quantitative methods, but is not a network; in fact I showed the limitations of traceable networks in this case for capturing the flow of information.

In the final chapter, I will consolidate some of the lessons learned about studying public science controversies on participatory platforms and think through some ways in which these visualisations can move forward from this point.

VII. CONCLUSION

When I started this project, I assumed that the wealth of data made available by new online platforms and digital tools built on top of them would in some sense deliver the project to me, making the qualitative work I was more familiar with easier, extending my reach. The idea was to bring the exciting new approaches developed in Mapping Controversies and Digital Methods into dialogue with the more qualitative, ethnographic tradition of ANT-informed controversy analysis, which they emerged from.

Yet the object of controversies, when seriously attended to – not just as a social phenomenon but as an analytic category – precisely demands that researchers let go of such pretensions. It insists on a certain fidelity to the empirical at the expense of concepts, intellectual baggage, and it seems, even methods and tools. This unexpectedly led me to try out new, visual, ad hoc techniques of analysis and visualisation, which complicate the way we think about the relations between quantitative and qualitative approaches. Although I often got the sense that my project became waylaid and weighed down by too many questions of method, barely scratching the surface of the controversial object of nuclear power, the resulting techniques have wider implications than just another cross-platform study of public science controversies.

In this final chapter, I will sum up these main methodological gains and point to some directions for further work. I will also attempt to characterise in more precise terms the digital research techniques I have developed in terms of their 'double social life' (Law et al., 2011), and address the related questions of what sort of collectives they perform, what silences or methodological 'hinterland' do they stake out, how these techniques might be developed further.

First, however, I would like to reiterate the premise and main problematics of the thesis and the solutions I arrived at.

01. SUMMARY

The main starting point of this doctoral project was the observation that the new online platforms that STS scholars use to map controversies are not innocent 'intermediaries'. Controversy analysts should recognize that these platforms produce very partial representations (Gillespie, 2010; Rogers, 2000, 2013b) and, it seemed to me, it may even be the case that they can impact their potential settlement. I argued that this warranted a study of online platforms 'in action' not, it should be said, as a critique of mapping techniques but as a supplementary exercise to enrich and extend this work. However there are some barriers to realizing this objective, first and foremost the possible scepticism within STS towards what could be conceived of as trendy and superficial digital technologies. In the first chapter, I tried to show why online platforms should be taken seriously in relation to controversies, discussing the example of the Fukushima disaster, but this case brought into view some confusion surrounding the terms in which we understand the contribution of platforms to controversy: the advancing of knowledge claims; increased and enriched participation; circulating alternative articulations in the media?

I cautioned that studying these platforms 'in action', inspired by more ethnographic studies of techno-scientific objects and settings, raised several conceptual and methodological problems, common to both qualitative and quantitative approaches, which this thesis then aimed to address.

- How to delineate an online study when controversies transcend particular online platforms?
- How to define what is relevant to a given study when these platforms have their own relevance-defining metrics which not only measure but shape interactions?
- How to track dynamic information flows within or between platforms?

Some of these problems, I argued in Chapter II, could be traced back to a certain ambivalence within STS about not just online media technologies but the media generally. Although there is a growing body of work bridging STS and media, these studies did not readily supply an empirical and conceptual understanding of media or platforms approaching the understanding of science 'in action' as outlined by ANT. What is interesting about media as such is their capacity to distribute content widely,

but for various reasons both practical and conceptual, researchers from STS and media studies have relied on understandings of media which treat the contents and the infrastructure and in which media are studied one at a time, rather than as part of an ecology. I argued that an 'in action' study requires addressing these conceptual hang-ups.

My solution was to decentre the object of controversy analysis away from platforms. This entailed using the STS concept of devices – which proposes a more fragile, heterogeneous and distributed account of technologies – and then placing these devices in tension with particular controversies. Controversies, I argued, bring into relief the contingency of platforms and the interdependencies between different media. And yet it was also important to recognise that our access to these controversies is only made possible through these platforms in the first place. I argued in favour of maintaining a tension between instrumentalising platforms to track controversies and studying the platforms themselves.

However, one challenge, which becomes more acute when we try to analyse controversies with platforms is the existing division of labour between quantitative and qualitative techniques. While it has been argued that digital data provide the means to bridge quantitative and qualitative approaches, and this found some success with web data: platforms seem to complicate the fluid ontology this presumes. I offered reflexive Digital Method tools and a battery of qualitative techniques as a solution but I also insisted that following the object of controversies means questioning the centrality of these platforms and dominant platform-objects within the study and methods tied to them. Instead of quantitative mapping techniques being assigned to readily calculable data and qualitative methods assigned to everything else, it is important to question the way platform data structures distribute these approaches.

In each of the empirical case studies, through engagement with specific controversies and specific platforms, I focused on ways in which methods could be partially decoupled from the platforms being studied. These chapters each addressed one of the three complications mentioned above.

Methodological Gains

Chapter IV examined the coverage of the Fukushima disaster on Wikipedia, and attempted to address the indeterminate boundaries of online settings. On the one hand, platforms like Wikipedia through its system of hyperlinks and comprehensive archiving reveals connections between Wikipedia and offline settings. I showed that tracing the controversy (mediated by the device) means following links outside of the platform by qualitative means. Yet the structuring of its data may in some ways impede these sorts of analyses. Since the controversy, on Wikipedia at least, centred around the parsing of external references – I went against the available data of the device to chart the composition of these references over time revealing how wikipedians relied on self reporting by the nuclear industry, against their normal code. I also showed, through qualitative analysis, how these users in concert with bots and scripts resisted the 'official' account offered by TEPCO and the Japanese Government, while negotiating policy and technical requirements. This approach allowed me to study a situation in which important asymmetries of resources arise beyond boundaries enforce by platforms.

Chapter V addressed the influence of metrics and quantifiable data both on the research design and the phenomena being researched. I did so through an investigation of online and offline activist interventions in the Hinkley Point siting controversy. Because the activities of distributed social movements straddle quantitative and qualitative analyses and micro and macro scales, I proposed visualising both traces of interactions and textual content as bi-partite networks. However, the fact that quantitative traces such as likes and shares not only describe interactions, but elicit and shape textual content as well, suggests that they should be read *together*. The tri-partite graphs, which brought together numerical and textual data, demonstrated the mutual influence of numbers and content – how Facebook may encourage more calculative behaviour. I also noted asymmetries inherent in adopting such a quantitative understanding of political participation: activists will never be able to compete in terms of likes with corporate entities – but I also argued that this is not the only criterion with which to evaluate these interventions.

Finally, Chapter VI addressed the elusive question of how information spreads online. The temptation with social media platforms is to think of them as instruments of information diffusion, merely circulating content, but I argued that social media

content, too, is always changed in the process of spreading. These transformations arise out of the interrelations between content and infrastructure, and to observe this I argued that we first needed a stable data object to follow, in this case I chose hyperlinks, which can be examined in terms of how it becomes modulated as it spreads. Building on the previous two chapters, visualising less accessible data and bringing quantitative and qualitative traces closer together, I proposed a solution in which the quantitative and qualitative analysis were intertwined – analysing an individual tweet and viewing wider patterns was just a matter of zooming in or out. This allowed me to show that different patterns of sharing associated with certain content were consequential for how far particular links spread.

The above methodological and empirical gains, however, were only possible because I was investigating these platforms 'in action' – through specific controversies, the specificity of which sometimes led me to go 'against the grain' of primary platform data structures. Controversies also allowed me to move past either cultural or technical explanations of platforms but also sender-receiver models of media, revealing a tangled web of different media and largely invisible backchannel communications (such as RSS, bots and less public discussion threads). Although very provisional, and admittedly not particularly 'nice' to look at, each of the quali-quantitative visualisations I developed both describe shifts in a controversy and some of the politics of platforms which might have remained more obscure if I had relied on either computationally advanced techniques or small scale qualitative methods on their own. The secret was to first recognise that, due to tension between controversies and devices, what is most relevant to the study is not necessarily what is most easy to analyse.

02. IMPLICATIONS FOR THE LITERATURE

So what do these interventions mean for the growing STS practice of mapping controversies with digital methods? The initial premise of this research was that platforms do not only enable the mapping of controversies with digital tools, they may play a role in enacting them as well, and are thus are worth studying with the approach of ANT-informed controversy analysis. Although it is impossible to generalise from just a handful of case studies, the contribution of the platforms I focused on remained somewhat ambiguous. While I was, from the beginning, sceptical about the more utopian claims in the wider literature about the potentials of platforms, in English speaking Wikipedia and UK centred social media, there were seemingly few decisive

interventions or impacts on controversies on the order of citizen-made geo mapping platforms in the Fukushima disaster.

While I hesitate to make any claims about the potential of platforms in controversies generally, what this study has shown is that such appraisals depend very much on what frame of reference we approach platforms with. If we are expecting platforms to intervene in knowledge on a scientific plane, then we may well be disappointed. Similarly in terms of participation, these platforms help coordinate protests and accrue members but it was not clear what this added on top of existing web sites and email lists. However, if we understand platforms as implicated in publicising the controversy with respect to more traditional media, then there are more obvious contributions to be observed. This might however require not looking for closure in the sense of science controversies but moments of politicization or publicization in the sense of the formation of issues. This is what I intended by the concept of 'digital demonstrations', if we think of platforms in terms of making things visible, both advancing articulations of a controversy or materialising an opposition, then these platforms do have a contribution to make, as long as we do not accept a platform definition of what makes these interventions successful. Direct action, Twitter storms, symbolic gestures and turns of phrase, what I referred to at one point as framing, may not register in terms of platform metrics but they are legitimate and potentially important strategies none the less.

This leads to one of the more unexpected revelations of the project: I originally thought that one could analyse these platforms armed mainly with existing STS literature on the construction of facts and public participation, but actually in each case study, the media or a media frame of reference emerged as an important topic of inquiry. Platforms were inextricably if ambivalently linked to online news, critiquing, sharing, commenting or circumventing it, but at least in some sense dependent on its rhythms. While I am not arguing that scholars of controversy need to completely re-orient themselves to the media, they might need to understand how a news story breaks or other mechanisms of publicity and this may necessitate drawing on literature from media studies in order to make sense of these routines and cycles.

While the contribution of platforms remained elusive in these case studies, it was, however, quite clear that these platforms are not *neutral* (Gillespie, 2010): that power asymmetries are scripted into their code and culture and these have implications for

the representations produced, even though these imbalances are contingent and open to subversion. This will be intuitively accepted by most STS researchers: platforms are not intermediaries, they are mediators (Latour, 2005). Yet, it should also be intuitively accepted that any attempts to represent a controversy shape or format it in some way (Hilgartner, 2000). But demonstrating this empirically means that cartographers of controversies might need to interrogate seemingly inconsequential bots or scripts and technical features and customs of these platforms in their analyses.

So the lesson for Mapping Controversies approaches is to take more seriously the seemingly incidental and banal media effects of these devices. Now, to the extent that mapping controversies is predominantly a pedagogic exercise, then considerations of this kind may not be feasible when students are trying to get to grips with already quite 'hairy' objects, but the consideration of platform effects on controversies could pertain to more experienced researchers attempting this work, or perhaps researchers more accustomed to offline controversies looking to add in an online component. Alternatively, the study of platform effects in controversies could take place at the early more exploratory states of research as a way of identifying statements, key actors and institutions but also selecting between different sources of data and information about the controversy.

But while it is important to take seriously the specificity of media or platforms in controversies, controversies also have implications for recent approaches arising from within media studies, which focus on media technologies. These traditions in various ways adopt the ANT or STS focus on materiality but not always the approach of controversy analysis with which it is sometimes associated. The object of controversies problematises the boundaries of settings, or in this case, particular media or platforms, and complicates existing models of media revealing a complex interacting web of communications.

There are plenty of existing ways to conceptualise these interrelations between media or platforms (Bausinger, 1984; Fuller, 2005; van Dijck, 2013) and practical ways to study them (Helmond, 2015) but my proposal is that controversies offer an especially effective way to bring these ecologies or ensembles into relief. They have the further benefit of demonstrably problematizing preconceptions of what counts as mainstream versus alternative, in Andrew Barry's example (Barry, 2001), or old versus new media,

and especially the presumed hierarchy of information flows, between scientists, journalists and audiences.

The techniques developed in this thesis can also hopefully compliment platform-specific or Digital Methods style studies by opening up less obvious or readily analysable devices, technologies and scripts for analysis and also revealing the complex interrelationships between platforms and less readily formatted websites. 'In action' studies could also be used to calibrate the standard battery of tools, helping to understand how bots, or news stories or press releases inflect the platform specific maps we use in specific cases.

Controversies and Devices

So following controversies can help locate new avenues for researching devices, but reflecting back on the case studies, it is clear that not all controversies are created equal. The Fukushima disaster was a 'hot' controversy *par excellence*: unplanned, extreme, redefining all that comes before as well as after it. Yet as the controversies unfolded, they tended to become weaker, more contrived affairs, creating less violent shockwaves in the platforms. In many cases, this was because the controversies were being actively de-politicised and managed by government and nuclear industry PR campaigns. These type of controversy logics may have implications for how much platforms can intervene, for example in uncertain disasters like Fukushima, but also have different affordances for empirical studies.

In some sense, I learned the most about Wikipedia because the controversy over the coverage of Fukushima threatened Wikipedia's *modus operandi* as an encyclopaedia. In contrast, controversies over the Hinkley new build project in the UK, manifested largely through protests and online interventions, were more like weak aftershocks of Fukushima and did not shake Facebook to its foundations, though it did reveal some tensions around Facebooks's possible censorship of political posts. In relation to Twitter I discussed a highly coordinated series of 'pseudo events', attempting to enrol (social) media actors and the relative *failure* of media actors to make these events controversial by connecting them to on-going events in Japan. Because this was, by then, a 'cold' controversy, I probably gained a bit less insight into Twitter as a device because it was not itself engulfed in the debate: even when Anonymous managed to perform a Twitter storm, online news media largely failed to recognise them. However,

one could also argue that the approach to Twitter analysis developed in this thesis is the most easily generalizable and translatable to other controversies or events, whereas the visualisation of Wikipedia references would mainly be pertinent for extreme controversies over news events, not every sort of article. So the intensities of controversies may have implications for what we can or cannot learn about particular devices.

Aside from the relative intensity or duration or stakes, the controversies I studied also varied in the extent to which they were dealt with along scientific or technical lines. For this reason, it is worth keeping open the question of whether or not we are dealing with techno-scientific controversies, in the sense that STS scholars are familiar with, or *issues* more generally. The later may involve more of an understanding of publicity and the challenges involved in gathering, focusing and sustaining attention and this might require different tools and approaches.

In the same way, it could be said that different platforms and devices have different capacities for research. Conveniently, the chapters on particular platforms were arranged in the order in which these platforms were created: Wikipedia (2001) Facebook (2004) Twitter (2006). In the conceptual lineage of devices identified by Ruppert and Law (2013) it could be said that Wikipedia becomes easier to understand through a more Foucauldian reading of device as 'apparatus' while the newer platform of Twitter is better understood as a rhizomatic, Deleuzian assemblage. But controversies tend to complicate these conceptual distinctions: Wikipedia ironically offered the most obvious avenues and tools for contesting the accounts of powerful actors and the circumventing of standard policy, while the supposedly more decentralised and 'revolutionary' social media platforms, at least in the few cases studied, reinforced asymmetries of resources between individual users and powerful, institutional actors.

So these are some of the implications for Mapping Controversies, Digital Methods and controversy analysis. In the penultimate section I will try to explore further the implications of these techniques for analysing controversies and digital social research more generally.

03. THE SOCIAL LIFE OF QUALI-QUANTI METHODS

Perhaps the most general contribution of this thesis, though somewhat incidental to the starting premise, was toward the development of broadly quali-quantitative approaches. This project has affirmed that digital data may in some ways make this project feasible. I have also widened the available arsenal of visualisations techniques beyond the trusted network diagram. Networks have worked so well in the past because of the way they map on to the object being studied: hyperlinks, online profiles and new forms of association in social media. The move of this thesis has been to shift the object away from these dominant devices, using particular controversies, and I showed that quali-quantitative methods are still feasible even with less formatted and structured data.

In this section I want to talk more specifically about the data visualisations I presented in this dissertation and how they could be pushed further, but also dwell on some of their politics and limitations. I first however want to talk about their ' double social life' (Law, et al 2011), in the sense of their relationship to existing techniques and approaches and their entanglements with platforms themselves.

As I explained, the Digital Methods approach is to stick closely to objects and methods already in the platforms, so that they can themselves be analysed as topics. However, they do also add something to existing traces through the imposition of (uncanny) social science methods like citation analysis and co-word (Marres and Gerlitz, 2015). The point I made in the methodology chapter, drawing on STS studies of representational practices in natural science, is that we might actually benefit from manipulating them *more*, instead of less, making less formatted data available for analysis and juxtaposing and combining data in new ways (Guggenheim, 2015). I scraped text using ad hoc scrapers, parsed it with Excel, squeezed text into spreadsheets and then back into text files.

However, these manipulations involve some violence to the original data which has costs and consequences in terms of what details or context may be lost or downplayed in the study. So I also want to contrast my practice of transformation and abstraction from that practiced in natural science and others in social science. I will do so with some *ex post facto* observations about what the visualisations have in common.

Exploratory Not Explanatory

Firstly, as sociologists of science have long noted, various types of visual and numerical inscriptions are used at every stage of the scientific process, and only some of these make it in to the final presentation of findings in scientific papers (Lynch, 1988). So there are images whose primary function is to convince and channel the reader through the argument as a rhetorical tactic in final reports, but there are also inscriptions which are used as *exploratory* devices in scientific practice. These are often part of chains of transformations (Latour, 1999) in which images are made increasingly abstract so that they can be better analysed and compared. So firstly I want to associate my practice with the more exploratory use, as an aide to producing textual representations rather than the presentation of findings in which the goal is to effectively communicate a particular set of claims.

In Mapping Controversies, the goal of data visualisations is explicitly to make the controversy more legible to stakeholders or publics and the visualisations. These are open and descriptive maps rather than explanations in the sense of statistical charts used to justify a point, but they are still in some sense prepared and cleaned for a particular audience. Digital Methods maps are in contrast not cleaned in the same way but they are still in some sense the final result of the research process as well as being used at intermediate stages. Although there is no need to draw a distinction too sharply, when visualisations are placed largely in the service of producing textual accounts they have different capacities as research devices in terms of how they handle complexity.

Maintaining Complexity

Venturini and his collaborators on the Electronic Maps to Assist Public Science (EMAPS) project have recently talked about the tension between complexity and usability in controversy maps (Venturini et al., 2015). Rather than aiming at a sort of equilibrium between the two, the authors describe the design process as a cyclical movement between the two poles as collaborators are enrolled to critique / enrich the maps. Clearly, the visualisations I have presented are very much on the complex end of the spectrum and the exploratory, analysis stage of the cycle. This is because I have in each case, following Digital Methods approaches, endeavoured to not edit the data after the fact, leaving room for surprises and accidents of the research process to occur.

While in each case I have purposefully selected certain digital traces and appropriate representational strategies to visualise them, I have wherever possible not reduced or abstracted the available data for the purposes of extracting nominal or ordinal categories or typologies. Lynch, discussed earlier (Lynch, 1988), describes a sliding scale in scientific visual practice from empirical representations to 'eidic' representations, which more or less gesture toward pure forms. We might associate the eidic with the regular social science practice of operationalizing social structures, categories or shared attributes in survey data – (class, race, social capital). I did not group certain data, seen as related to normative categories in order to make causal claims. When I did group typologies of data, these followed categories divined by research subjects/platforms – domains in the Wikipedia graph or identical text in the Twitter graph.

One of the claims about the advantages of transactional or by-product data is that we can to some extent deal with populations not samples (Savage et al., 2010) so that this reduction is not necessary. While this was technically speaking possible in the case of Wikipedia's references – every reference in every version of Wikipedia was available as a data point – yet due to limitations of computing power, I was forced to sample the references 1 in every 10 and then every 100 versions. In the Facebook chapter however, I could display every single post and every single user, though as I will describe below, the terms in the graph were automatically reduced by the Alchemy database. As mentioned earlier ANTA's co-word process involves automatically extracting key terms and merging different spellings and namings into single entities. This should be avoided if possible, because different spelling and non-proper names can still be important for the study, but it is to some extent necessary from a computing standpoint to at least use stop lists to weed out common words like ('and', 'the', 'a', etc.). In the case of the Twitter URL sequencer, while the data set was circumscribed in relation to shares of a URL, all of the tweets were included in the visualisation. Tweets were never removed based on given criteria, only made more or less available to the researcher's wandering eye through selective highlighting.

The main exception is that each of the visualisations above were circumscribed by a particular window of time. This was often arrived at by starting with frequency measures to identify potential controversies and sub controversies and giving a relatively wide birth to allow them to fully unfold. This leads to another affordance of digital data, the ability to use continuous time, rather than relying on time slices

(Savage et al., 2010). Wherever possible in these visualisations, I have endeavoured to use more granular displays of time often using time stamps, which are precise to the second. Thus the slices of Wikipedia references and the Facebook posts reveal certain rhythms and concentrations of activity, which would not otherwise be visible. Absolute time was not possible for the Twitter URL visualisation – the tweets are currently placed in time order but not spaced according to time stamp – something which could be implemented in future. One can get a sense of the rhythm of posting from the frequency graph of the URL shares I produced however.

When visualisations are primarily in the service of producing textual accounts, freed from certain requirements for easy communications, there is not necessarily such a trade off between complexity and legibility. The trick is to find ways of displaying the data (through colour coding, spacing etc.) which allow our eyes and minds, rather than algorithms and scripts, to identify relevant patterns. This may require training and familiarity, like a doctor looking for hairline fractures in an x-ray, but these graphs are not illegible *per se*, they just require different ways of seeing.

Revealing the Conditions of Production

The third way in which my graphs are different from other types of related representations is the way they actively display rather than conceal artefacts of their production. Earlier I used Helen Verran's work (2012) to talk about the performative effects of numbers and the importance of understanding their indexicality, or relations with the wider world. In one sense, revealing the conditions of production is a way of keeping the numbers relation to the entanglements they are necessarily extracted from, including their categorical and value laden production. In other words, they can be made somewhat reversible. The same could be said however of other sorts of discreet data points and their embedding in larger socio-technical devices.

In the Wikipedia graph, this is mainly in relation to the 'No Link' strip. As mentioned earlier, my rudimentary scraper worked to varying degrees because some of the reference links were placeholders, made by bots, mistakes or deployed casually without a thought of other users (or social scientists). So in order to draw attention to this I provided a category called 'No Link' in the diagram, which eluded to numbered references without hyperlinks. This keeps in the picture both deviations in the users behaviour in relation to references *and* deviations in the scraper's ability to capture

data. In relation to Facebook, rather than abstracting social relationships or semantic clusters from interactions on a page I decided to keep the posts, the original source in the diagram through bi-partite graphs. The users were not deleted or changed but the words (again unformatted and harder to grasp) were again transformed through the Alchemy database of proper names. However, in the ANTA interface, I avoided removing any further words so that certain accidents of the selection process such as the appearance of posts like 'Daily Nuclear Update' could become visible. On Twitter, when the different shares of the URL were parsed into several basic types, the original tweets were kept on the left hand of the graph so that errors in the process could be easily detected (when two unrelated tweets were deemed to be identical or when identical tweets were misrecognized).

Retaining some of the traces of the maps' production in the visualisation, aside from checks and balances in the research process, is one of the ways that data visualisations can interface better with qualitative techniques by giving researchers threads to follow. These accidents of the process interrupt the smooth reading of graphs and force the analyst to read the graphs as constructed and opaque data rather than transparent sources of information.

Claims and Automation

I have proposed that in situations of data deluge and radical uncertainty, it may make sense to use data visualisations as interpretive tools for producing textual accounts, rather than as end products in themselves – either for generating more numbers like p-values and odds ratios or for communicating findings. However, there is nothing inherent in these approaches, which precludes the development of generalisations or normative claims or push button analyses at a later stage.

Especially in the previous chapter, I started making some normative claims about 'grassroots', 'broadcast' and 'spin' styles of link sharing. I hesitated to do so, being aware that the standard practice in ANT is to provide a minimal 'infralanguage' and otherwise stick as closely as possible to the terminology actors themselves use to make sense of the world. Twitter does not yet have terms for these behaviours but these categories follow closely concrete practices (using @'s, deploying RSS bots, discursive texts), which could be easily operationalized. These could be used as metrics to roughly understand or profile how links are being shared. Similarly, claims could be made

about the relative heterogeneity or stability of Wikipedia references over time as an indicator of controversy, possibly as a supplement to analyses like Contropedia (Borra et al., 2014).

These are all interesting possibilities, which might make these STS studies more relevant to more traditional sociological, and media studies audiences or even participants and stakeholders (see below). The challenge however, is to not build on top of these observations further and further layers of mathematical operations: black boxing the original assumptions or empirical details they emerge out of. One would always need to be able to re-open push button tools in the light of unstable events and controversies.

The Politics of Decentring Devices

But why is it important that we search out and analyse less obvious, less readily formatted data – such as Wikipedia's references, unformatted text, email lists, or truncated URLs? On one hand there is perhaps a danger that we limit ourselves to the most popular or successful content, or adopt a view of the platform favoured by its architects or dominant users. This is fine so long as the design and systems of rankings themselves becomes the object of research, in the sense proposed by Digital Methods. But this might still be subject to the common criticism of ANT-inspired approaches, articulated best by Susan Leigh Star (Star, 1990) that ANT-researchers often study the networks of powerful scientists at the expense of marginalized actors who necessarily fall outside these networks (invisible labour etc.):

> ..the political order described in actor network theory, or in descriptions of the creation of scientific facts, they describe an order which is warlike, competitive, and biased toward the point of view of the victors (or the management). (Star, 1990)

So this is one reason why it makes sense to juxtapose the methods of the medium with qualitative approaches and with more relational, non frequency based measures or visualisations which highlight less obvious data, because these methods perform and make visible collectives, such as ephemeral publics and small groups of anti-nuclear activists, whose voices are not always heard against the din of social media. Searching out less obvious data, then has, potentially a political angle, revealing some of the power asymmetries in what becomes visible and what becomes obscured by devices.

These techniques accomplish this without debunking or critiquing the representations and collectives favoured by dominant data structures, but by widening our vantage and complicating these views. In general, this becomes possible when we define better what our research object is, in contrast to the definition offered by digital devices, by introducing controversy and contingency into the proceedings.

The Hinterland

As John Law describes, methods create the objects they are supposed to describe but they will also necessarily create absences and silences and boundaries, leaving a 'hinterland' of what remains unsayable (Law, 2004). One of the implications of this is that, to the extent that I used mixed methods and tools, these different types of representations did not converge on a shared reality lying underneath but produced multiple ontologies (Mol, 1999). So when we rely on different platforms to give perspectives on a controversy, these are each in their own way partial accounts. Even the jump between analysing maps and qualitative analysis of texts creates disjuncture. In the Wikipedia chapter, in which the references were mapped using different data than the discussions of the references, these created conflicting impressions of how democratic and subversive Wikipedia could be. The automated maps showed a bias toward nuclear industry and government sources but the textual analysis revealed creative manoeuvring behind the scenes. In contrast, the Twitter tool makes the act of reading and the visualisation parallel and compatible, allowing a more seamless ontology.

There are always procedures for choosing between or merging ontologies, as Anne-Marie Mol (2002) describes in relation to methods in medicine: while there are multiple representations of a condition (thus multiple bodies), surgery in her particular example, is the 'gold standard' which supersedes the others. However it is unclear in online research what the 'gold standard' is anymore? In the case of distributed controversies, it is no longer given that offline fieldwork or if individual follow up interviews will offer a definitive account of the controversy. As I suggested, offline sites and individual users can certainly add complexity to the analysis but are not themselves the proper defining unit.

Although I have tried to push computational tools further into this hinterland to accommodate different types of data and particularly different ways of assisting the

reading of texts, there are still many challenges. Firstly, images, which have proved so potent in cultural accounts of nuclear power (Weart, 2012) are very difficult to analyse computationally except in very superficial ways. The same goes for videos and sound. Manovich has pioneered ways of extracting formal properties out of images and displaying them in ways which allow for interpretive analyses (Manovich et al., 2011) and Michael Guggenheim's (Guggenheim, 2015 forthcoming) work eludes to ways of dealing with images which move beyond visual anthropology or sociology approaches.

It is however a truism of controversy analysis generally that no matter what the setting(s) being studied, there will always be some other setting: closed door policy meetings or economic institutions, or offline direct action which prove consequential for the controversy. This is unavoidable, there will always be a backstage to every seemingly backstage process (Schlecker and Hirsch, 2001), but I feel that increasingly even if decisions are made in smoke filled rooms they must be aired in public, potentially online. Like the strike price deal between EDF and the British government: media articulations are required to package and justify these arrangements and we can study this process of publicity even if these other sorts of processes remain black boxed.

Visibility and Ethics

Methods will always perform gaps and silences but there are also, conversely, problems with what they shine a light on. I earlier described how platform data structures, surveillance technologies and human actors make certain phenomena variably visible to both researchers and other users. I also described how our methods, which are largely linked to these devices might need to make other sorts of materials visible. But there is a danger in this, perhaps more active and intrusive use of methods and tools. Ethics in the age of big data is a somewhat uncharted frontier, especially at a time when Facebook has itself starting to do social psychological research with little consideration for ethical protocols.[202] While I cannot offer a definitive or programmatic statement on the matter, I will try to distance my approach from two positions which I think are less helpful in this sort of research.

[202] As discussed recently in *The Atlantic*: http://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/ (Accessed 10 September 2015)

Firstly, one position might hold that we have already produced research ethics for the internet, such as the popular AoIR guidelines[203] and new online platforms, in particular those called 'social media' add nothing new to this. The first wave of internet ethics arose at a time when the anonymity of users was key, and arguments were made about the strong attachments people made to these avatars and virtual relationships (Markham, 2005). Privacy and the need for informed consent were central but often permission was understood from a human subjects model of ethics in natural science and social psychology (Bassett and O'Riordan, 2002). But today, platforms like Facebook demand that its users maintain a stable identity and are knowingly 'open' and 'public'. It is nearly impossible to effectively conceal identities in an age of Google search, where direct quotes and contextual information is easy to locate. Obviously, we need to protect information in closed groups and negotiate access but I think social media does fundamentally change what constitutes private and public content.

However, the other erroneous way of proceeding is what we might call the 'terms and conditions' version of ethics. This would be to align our ethics to the legal position of the platforms. If users agree to the terms and conditions, and understand that their data can be viewed and analysed, there is no problem (Zimmer, 2010). Elm (2009) describes how the need to obtain consent is normally premised on a public / private distinction which does not easily hold in certain spaces of the internet. I think it is more helpful, instead of presuming these categorisations of spaces to think instead of the sliding scale of visibility, made variously possible through monitoring technologies, algorithms and user behaviours – there is no ultimate vantage there are only ways of searching – and there are barriers and gradients of access, some of which are erected by research subjects and devices.

Wikipedia is by most understandings, 'public' and mostly anonymous, but in the course of wider investigations, I found certain sections of the website to be more backstage and frequently found users behaving as if they were in a private conversation, in less obvious forums such as user pages and topic based discussion boards. The same went for the ostensibly public spaces of Twitter and Facebook: while any content can potentially be viewable, I found that, following Murthy (2013), statements had an implied perceptual range. Not every utterance is expected to reach millions, they are premised on a presumed audience and I have deliberately avoided publishing certain

[203] Available from: http://aoir.org/reports/ethics.pdf (Accessed 15 September 2015)

exchanges which seemed to be less publicity facing. Other users, perhaps from an older generation may not understand fully the consequences of Tweeting. But methods of social research are also tools of visibility and it is possible for researchers to shine a spotlight on these darker corners of the web. One can for example draw attention to actions which were not obvious at the time and 'out' users even when their actions are already, technically, public. As with everything else in this thesis, ethics need to arrive from the empirical rather than be imposed from above, so the damaging side effects of research need to be carefully considered on a case-by-case basis.

It helps, in general however, to study actors like experienced activists and PR professionals, who are specifically publicity-facing and self-promotional, who want their message disseminated as widely as possible. Following the actors means that researchers can better understand these different levels of visibility and publicness, but researchers also need to consider how (human) research subjects participate in research as well.

Participation

In one important sense described by STS, research subjects participate in research to the extent that researchers are following their attempts to make sense of a controversy. Actors participate in setting the boundaries and defining what is relevant to the study. Yet they are not necessarily aware of this role. One way to resolve certain questions of access and ethics is to make research more interactive from the beginning, to enrol stakeholders in the controversy with a view to not just studying the controversy but looking for resolutions to it (Venturini, 2010a). There have also been proposals for methods which more obviously intervene in the object being studied (Wilkie et al., 2015). My feeling is that in the case of new phenomena, such as participatory media, there is actually a shortage of simple sociological descriptions, which need to be produced along side these other projects.

However, I propose that researchers can also bring research subjects into the process in a more traditional way, though methods like interviews and participant observation. The case studies in this thesis have produced lots of questions, lists of important actors and offline settings; so the next step is to build these back into the methodology. Instead of starting with the offline site to determine the scope and direction of online analyses – here the online can be used to structure the offline work.

Although there is not space to discuss this, I have already started to implement this qualitative work: attending protests and interviewing activists and policy experts. These interviews can problematize or feed back into the online analyses – generating new lists of texts, actors, events venues and texts. There is also the possibility of using data visualisations as a kind of 'photo-elicitation' method: showing visualisations to users and pointing out their position in them can help reflect on the case but also the process of map making and what it leaves out.

This however evokes a previous attempt by a sociologist (Touraine 1982) to share findings with, appropriately enough, anti-nuclear activists. Touraine, according to later accounts, came to the situation with an idea of what the nuclear activists *should* be doing and this came out in his field notes – angering many of the participants. One of the goals of producing descriptions as opposed to explanations is that participants are more able to make their own meanings out of this. However, as STS studies of participation have shown, there will always be unavoidable asymmetries in the terms by which participation happens, even between researchers and research subjects.

04. CONCLUDING REMARKS

Although this thesis remained firmly in the realm of STS and the intersection of media and the cartography of controversies, these studies do have implications for research on the internet or big data analyses more generally. The problem with much of the data visualisations and automated tools, which emerge not just from social research but also tech companies themselves, is that they have a seductive rhetorical power (Kennedy et al., 2014): to the casual viewer they are often black boxed, yet I think that they remain convincing because they are, to use Latour's term 'panoramas' (2006) they offer a falsely complete picture of some landscape of socio-technical life. They do not draw attention to their conditions of production; they smooth over the gaps and have cleaned the 'outliers'.

By partially disengaging these data visualisations from the platforms and devices they are normally associated with, these visualisations become very contingent *oligopticons*, methods which only present a partial slice of social life and resist easy analysis. This leads to another way of thinking about data visualisations – as media themselves.

Thinking back to Latour's concept of immutable mobiles: representations which travel easily and survive the journey intact (Latour, 1986): the quintessential representation he has in mind are numbers because they travel effortlessly and are hard to change. The visualisations I have made could have been communicated in numbers, easy to understand bar charts etc. whose meanings are (relatively) unambiguous and easier to understand, more mobile. These would travel further and faster and could make it to wider audiences, and perhaps this will be necessary to compete with push button big data analyses which claim to require no contextual or qualitative understanding at all.

While this thesis argues that sociologists should embrace the use of platforms and digital tools, they might also benefit from questioning the way both interfere and mediate research. Just as ANT recommends tracing the social through small dirt roads as opposed to superhighways, we need visual techniques which force the user to stop and explore, which stubbornly refuse reductive or monolithic readings. While I started the project thinking that platforms and digital tools would make qualitative analysis easier I have come to the conclusion that while digital tools do enhance research possibilities, if used in a concerted way, they might make research slower, and in an age of 'real time' (Marres and Weltevrede, 2013) research, this may not be such a bad thing.

REFERENCES

Allan S (2006) *Online News: Journalism and the Internet*. Maidenhead: McGraw-Hill International.

Amir S (2009) Challenging nuclear: antinuclear movements in postauthoritarian Indonesia. *East Asian Science, Technology and Society* 3(2-3): 343–366.

Amit V (2000) *Constructing the Field: Ethnographic Fieldwork in the Contemporary World*. London: Routledge.

Anderson B (1991) *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.

Ang I (1991) *Desperately Seeking the Audience*. London: Routledge.

Ang I (1992) Living-room wars: new technologies, audience measurement and the tactics of television consumption. In: Silverstone R and Hirsch E (eds), *Consuming Technologies: Media and Information in Domestic Spaces*, Hove, UK: Psychology Press, pp. 74–81.

Aubrey C (1991) *Meltdown: The Collapse of the Nuclear Dream.* London: Collins and Brown.

Barnes B (1977) *Interests and the Growth of Knowledge*. Milton Park: Routledge.

Barry A (2001) *Political Machines: Governing a Technological Society*. London: Athelone Press.

Barry A (2005) The anti-political economy. In: Slater D and Barry A (eds), *The Technological Economy*, Taylor & Francis, pp. 84–100.

Barry A (2010) Tarde's method: between statistics and experimentation. In: Candea M (ed.), *The Social After Gabriel Tarde: Debates and Assessments*, Abingdon: Routledge, pp. 177–190.

Barry A (2012) Political situations: knowledge controversies in transnational governance. *Critical Policy Studies* 6(3): 324–336.

Barry A (2013) *Material Politics: Disputes Along the Pipeline*. Malden, MA: John Wiley & Sons.

Bassett EH and O'Riordan K (2002) Ethics of Internet Research: Contesting the Human Subjects Research Model. *Ethics and Information Technology* 4(3): 233–247.

Bastian M, Heymann S, Jacomy M, et al. (2009) Gephi: an open source software for exploring and manipulating networks. In: *Proceedings of 3rd International Conference on Weblogs and Social Media (ICWSM)*, Menlo Park, CA: AAAI, pp. 361–362. Available from: http://www.aaai.org/ocs/index.php/ICWSM/09/paper/viewPDFInterstitial/154Forum/1009 (accessed 19 August 2015).

Bateson G (2000) *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Chicago: University of Chicago Press.

Bausinger H (1984) Media, technology and daily life. *Media, Culture & Society* 6(4): 343–351.

Beaulieu A (2005) Sociable hyperlinks: an ethnographic approach to connectivity. In: Hine C (ed.), *Virtual Methods: Issues in Social Research on the Internet*, London: Berg, pp. 183–198.

Beck U (1992) *Risk Society: Towards a New Modernity*. London: Sage.

Beck U (2000) Risk society revisited: theory, politics and research programmes. In: Adam B, Beck U, and Van Loon J (eds), *The Risk Society and Beyond: Critical Issues for Social Theory*, pp. 211–229.

Beer D (2009) Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society* 11(6): 985–1002.

Benford RD and Snow DA (2000) Framing processes and social movements: an overview and assessment. *Annual Review of Sociology* 26(1): 611–639.

Benkler Y (2006) *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. London: Yale University Press.

Bennett L and Segerberg A (2012) The logic of connective action. *Information, Communication & Society* 15(5): 739–768.

Bennett WL, Segerberg A and Walker S (2014) Organization in the Crowd: Peer Production in Large-Scale Networked Protests. *Information, Communication & Society* 17(2): 232–260.

Blackmore S (2000) *The Meme Machine*. Oxford: Oxford University Press.

Bloor D (1976) *Knowledge and Social Imagery*. Chicago: University of Chicago Press.

Bloor D (1982) Durkheim and Mauss revisited: classification and the sociology of knowledge. *Studies in History and Philosophy of Science London*, Special Issue: The Sociology of Knowledge: Cases and Debates. 13(4): 267–297.

Boczkowski P and Lievrouw LA (2008) Bridging STS and communication studies: scholarship on information technologies. In: Hackett EJ, Amsterdamska O, Lynch M, et al. (eds), *The Handbook of Science and Technology Studies*, The MIT Press, pp. 949–977.

Boczkowski PJ (2005) *Digitizing the News: Innovation in Online Newspapers*. Cambridge, MA: The MIT Press.

Boczkowski PJ and Siles I (2014) Step toward cosmopolitanism in the study of media technologies: integrating scholarship on production, consumption, materiality and content. In: Gillespie T, Boczkowski PJ, and Foot KA (eds), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, MA: MIT Press, pp. 53–76.

Boorstin D (1961) *The Image: A guide to pseudo-events in America*.

Borch C (2012) *The Politics of Crowds: An Alternative History of Sociology*. Cambridge, UK: Cambridge University Press.

Borra E and Rieder B (2014) Programmed method: developing a toolset for capturing and analyzing tweets. *Aslib Journal of Information Management* 66(3): 262–278.

Borra E, Weltevrede E, Ciuccarelli P, et al. (2014) Contropedia-the analysis and visualization of controversies in wikipedia articles. In: *Proceedings of The International Symposium on Open Collaboration*, New York: ACM, p. 34.

Bowker G and Star SL (1999) *Sorting Things Out*. MIT Press.

boyd danah (2009) A Response to Christine Hine. In: Markham AN and Baym NK (eds), *Internet Inquiry: Conversations About Method*, Thousand Oaks: Sage, pp. 26–32.

boyd danah and Crawford K (2011) Six Provocations for Big Data. *A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*.

boyd danah and Crawford K (2012) Critical questions for big data. *Information, Communication & Society* 15(5): 662–679.

boyd danah, Golder S and Lotan G (2010) Tweet, tweet, retweet: conversational aspects of retweeting on Twitter. In: *HICSS-43*, Honolulu, HI: IEEE, pp. 1–10.

Bruns A (2005) *Gatewatching: Collaborative Online News Production*. New York: Peter Lang.

Bruns A (2006) Wikinews: the next generation of online news? *Scan Journal* 3(1).

Bruns A (2007) Methodologies for mapping the political blogosphere: an exploration using the issuecrawler research tool. *First Monday* 12(5).

Bruns A (2009) From Prosumer to Produser: Understanding User-Led Content Creation. In: *Transforming Audiences 2009*, London: Produsage.org.

Bruns A (2012) How long is a tweet? Mapping dynamic conversation network on Twitter using Gawk and Gephi. *Information, Communication & Society* 15(9): 1323–1351.

Bruns A and Burgess J (2012) Researching news discussion on Twitter. *Journalism Studies* 13(5-6): 801–814.

Bucher T (2012) Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society* 14(7): 1164–1180.

Callon M (1980) Struggles and negotiations to define what is problematic and what is not. In: Knorr KD, Krohn R, and Whitley R (eds), *The Social Process of Scientific Investigation*, Sociology of the Sciences A Yearbook, Netherlands: Springer, pp. 197–219.

Callon M (1998) *The Laws of the Markets*. Oxford: Blackwell Publishing.

Callon M, Law J and Rip A (1986) In: *Mapping the Dynamics of Science and Technology*, London: Springer.

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

Castells M (2013) *Networks of Outrage and Hope: Social Movements in the Internet Age*. Cambridge, UK: Polity.

Castillo C, Mendoza M and Poblete B (2011) Information credibility on twitter. In: *Proceedings of the 20th International Conference on World Wide Web*, New York: ACM, pp. 675–684.

Cetina KK (2009) The synthetic situation: interactionism for a global world. *Symbolic Interaction* 32(1): 61–87.

Chateauraynaud F (2009) Public controversies and the Pragmatics of Protest.

Coleman G (2011) Hacker politics and publics. *Public Culture* 23(3 65): 511–516.

Collins HM (1987) Certainty and the public understanding of science: science on television. *Social Studies of Science* 17(4): 689–713.

Collins HM and Pinch T (1998) *The Golem: What You Should Know About Science*. Cambride, UK: Cambridge University Press.

Corner J, Richardson K and Fenton N (1990) *Nuclear Reactions: Form and Response in 'Public Issue' Television*. New Barnet, UK: John Libbey.

Couldry N (2008) Actor network theory and media: do they connect and on what terms? In: Hepp A, Krotz F, Moores S, et al. (eds), *Connectivity, Networks and Flows: Conceptualizing Contemporary Communications*, Cresskill, NJ, USA: Hampton Press, Inc., pp. 93–110.

Currie M (2012) The feminist critique: mapping controversy in Wikipedia. In: Berry D (ed.), *Understanding Digital Humanities*, Houndmills: Palgrave Macmillan, pp. 224–248.

D'Angelo P (2002) News framing as a multiparadigmatic research program: a response to entman. *Journal of Communication* 52(4): 870–888.

Danowski JA (2009) Inferences from word networks in messages. In: Krippendorff KH and Bock MA (eds), *The Content Analysis Reader*, London: Sage, pp. 421–429.

Deacon and Fenton N (1999) The natural history of a news item. *Media, Culture & Society* 21: 5–31.

Didier E (2010) Gabriel tarde and statistical movement. In: Candea M (ed.), *The Social after Gabriel Tarde: Debates and Assessments*, pp. 163–176.

Diminescu D, Bourgeois M, Renault M, et al. (2011) Digital diasporas atlas exploration and cartography of diasporas in digital networks. In: *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM'11)*, AAAI: Menlo Park, CA.

Doan S, Vo B-KH and Collier N (2012) An Analysis of Twitter Messages in the 2011 Tohoku Earthquake. In: Kostkova P, Szomszor M, and Fowler D (eds), *Electronic Healthcare*, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Berlin: Springer, pp. 58–66.

Dodge M (2005) The role of maps in virtual research methods. In: Hine C (ed.), *Virtual Methods: Issues in Social Research on the Internet*, Oxford: Berg Publishers, pp. 113–127.

Dorfman P, Fucic A and Thomas S (2012) *18 late lessons from Chernobyl, early warnings from Fukushima*. Late Lessons from Early Warnings: Science, precaution, innovation, Vol. 2, Copenhagen: European Environment Agency.

Driscoll K and Walker S (2014) Big Data, Big Questions| Working Within a Black Box: Transparency in the Collection and Production of Big Twitter Data. *International Journal of Communication* 8(0): 20.

Durant JR, Evans GA and Thomas GP (1989) The public understanding of science. *Nature* 340(6228): 11–14.

Elmer G (2006) Re-tooling the network parsing the links and codes of the web world. *Convergence: The International Journal of Research into New Media Technologies* 12(1): 9–19.

Elm MS (2009) How Do Various Notion of Privacy Influence Decisions in Qualitative Research? In: Markham AN and Baym NK (eds), *Internet Inquiry: Conversations About Method*, Thousand Oaks: Sage, pp. 69–87.

Entman RM and Rojecki A (1993) Freezing out the public: elite and media framing of the U.S. anti-nuclear movement. *Political Communication* 10(2): 155–173.

Entwistle J and Slater D (2014) Reassembling the cultural. *Journal of Cultural Economy* 7(2): 161–177.

Espeland WN and Sauder M (2007) Rankings and reactivity: how public measures recreate social worlds. *American Journal of Sociology* 113(1): 1–40.

Fairclough N (1995) Media discourse.

Farías I (2014) Misrecognizing tsunamis: ontological politics and cosmopolitical challenges in early warning systems. *The Sociological Review* 62(S1): 61–87.

Fielding JL and Fielding NG (2008) Synergy and synthesis: Integrating qualitative and quantitative data. Available from: http://epubs.surrey.ac.uk/231711/2/SRI_deposit_agreement.pdf (accessed 19 May 2013).

Freudenberg WR and Rosa EA (1984) *Public Reactions to Nuclear Power*. Boulder, Colorado: Westview.

Friedman SM (2011) Three Mile Island, Chernobyl, and Fukushima: an analysis of traditional and new media coverage of nuclear accidents and radiation. *Bulletin of the Atomic Scientists* 67(5): 55–65.

Fuller M (2005) *Media ecologies: Materialist energies in art and technoculture*. MIT Press.

Fuller M (2008) *Software Studies: A Lexicon*. Cambridge, MA: MIT Press.

Gamson WA and Modigliani A (1989) Media discourse and public opinion on nuclear power: a constructionist approach. *American Journal of Sociology* 95(1): 1–37.

Gans HJ (1979) *Deciding What's News: A Study of CBS Evening News, NBC Nightly News, Newsweek, and Time*. Chicago: Northwestern University Press.

Gerbaudo P (2012) *Tweets and the Streets: Social Media and Contemporary Activism*. New York: Pluto Press.

Gerbaudo P (2014) The persistence of collectivity in digital protest. *Information, Communication & Society* 17(2): 264–268.

Gerlitz C and Helmond A (2013) The Like economy: Social buttons and the data-intensive web. *New Media & Society*.

Gerlitz C and Lury C (2014) Social media and self-evaluating assemblages: on numbers, orderings and values. *Distinktion: Scandinavian Journal of Social Theory* 15(2): 174–188.

Gerlitz C and Stevenson M (2009) The place of issues. Digital Methods Initiative. Available from: https://wiki.digitalmethods.net/Dmi/ThePlaceOfIssues (accessed 6 July 2015).

Gieryn TF (1983) Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review* 48(6): 781–795.

Gieryn TF (1999) *Cultural Boundaries of Science: Credibility on the Line*. Chicago: University of Chicago Press.

Gillespie T (2010) The politics of 'platforms'. *New Media & Society* 12(3): 347–364.

Gillespie T (2014) The relevance of algorithms. In: Gillespie T, Boczkowski PJ, and Foot KA (eds), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, MA: MIT Press, pp. 167–194.

Gillespie T, Boczkowski PJ and Foot KA (2014) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press.

Gitlin T (1980) *The Whole World Is Watching: Mass Media in the Making & Unmaking of the New Left*. London: University of California Press.

Gladwell M (2010) Small Change. *The New Yorker*. Available from: http://www.newyorker.com/reporting/2010/10/04/101004fa_fact_gladwell?currentPage=all (accessed 24 February 2014).

Glaser BG and Strauss AL (1967) *The discovery of grounded theory: Strategies for qualitative research*. Aldine de Gruyter.

Glasgow University Media Group (1976) *Bad News*. Abingdon, UK: Routledge.

Goffman E (1974) *Frame Analysis*. Cambridge, MA: Harvard University Press.

Goffman E (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Gregory J and Miller S (1998) *Science in Public*. Cambridge, MA: Basic Books.

Guggenheim M (2015) The media of sociology: tight or loose translations? *British Journal of Sociology* 66(2): 345–372.

Guimaraes C (2005) Doing anthropology in cyberspace: fieldwork boundaries and social environments. In: Hine C (ed.), *Virtual methods: Issues in social research on the Internet*, Oxford, UK: Berg Publishers.

Hall S, Critcher C, Jefferson T, et al. (1978) *Policing the Crisis: Mugging, the State, and Law and Order*. London: Macmillan.

Hall S, Hobson D, Lowe A, et al. (1980) Encoding / decoding. New edition. In: Hall S (ed.), *Culture, Media, Language: Working Papers in Cultural Studies, 1972-79*, London: Routledge, pp. 117–127.

Hammersley M (1992) Deconstructing the qualitative-quantitative divide. In: Brannen J (ed.), *Mixing Methods: Qualitative and Quantitative Research*, Aldershot: Avebury. Available from: http://oro.open.ac.uk/20445/ (accessed 27 January 2015).

Hara N and Doney J (2015) Social construction of knowledge in wikipedia. *First Monday* 20(6).

Haraway D (1988) Situated knowledges: the science question in feminism and the privilege of partial perspective. *Feminist studies* 14(3): 575–599.

Hecht G (2009) *The Radiance of France: Nuclear Power and National Identity after World War II*. Cambridge, MA: MIT Press.

Hellsten I, Dawson J and Leydesdorff L (2010) Implicit media frames: automated analysis of public debate on artificial sweeteners. *Public Understanding of Science* 19(5): 590–608.

Helmond A (2015) The Web as Platform: Data Flows in Social Media. Thesis, Amsterdam: University von Amsterdam.

Hermida A (2010) Twittering the news. *Journalism Practice* 4(3): 297–308.

Hilgartner S (1990) The dominant view of popularization: conceptual problems, political uses. *Social Studies of Science* 20(3): 519–539.

Hilgartner S (2000) *Science on Stage: Expert Advice as Public Drama*. Stanford University Press.

Hilgartner S (2007) Overflow and containment in the aftermath of disaster. *Social Studies of Science* 37(1): 153–158.

Hilgartner S, Bell RC and O'Connor R (1983) *Nukespeak: The Selling of Nuclear Technology in America*. New York: Penguin Books.

Hine C (2000) *Virtual Ethnography*. Thousand Oaks, CA: Sage.

Hine C (2005) *Virtual methods: Issues in social research on the Internet*. Berg Publishers.

Hogan B (2010) The presentation of self in the age of social media: distinguishing performances and exhibitions online. *Bulletin of Science, Technology & Society* 30: 377–386.

Hogan B, Fielding N, Lee RM, et al. (2008) Analyzing Social Networks. In: *The Sage Handbook of Online Research Methods*, Thousand Oaks, CA: Sage, p. 141.

Howard PN (2002) Network Ethnography and the Hypermedia Organization: New Media, New Organizations, New Methods. *New Media & Society* 4(4): 550–574.

Irwin A (2001) Constructing the scientific citizen: science and democracy in the biosciences. *Public Understanding of Science* 10(1): 1–18.

Irwin A and Michael M (2003) *Science, Social Theory and Public Knowledge*. Maidenhead, UK: Open University Press.

Jacomy M, Venturini T, Heymann S, et al. (2014) Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLoS ONE* 9(6). Available from: v (accessed 19 August 2015).

Jasanoff S (1994) *Learning from disaster: risk management after Bhopal*. University of Pennsylvania Press.

Jasanoff S (2004) Ordering knowledge, ordering society. In: Jasanoff S (ed.), *States of Knowledge: The Co-Production of Science and the Social Order*, pp. 43–68.

Jasanoff S (2005) *Designs on Nature: Science and Democracy in Europe and the United States*. Princeton, NJ: Princeton University Press.

Johnstone P (2014) Planning reform, rescaling, and the construction of the postpolitical: the case of The Planning Act 2008 and nuclear power consultation in the UK. *Environment and Planning C: Government and Policy* 32(4): 697 – 713.

Katz E and Lazarsfeld PF (1955) *Personal influence: the part played by people in the flow of communications*. New York: Free Press.

Kelty C (2005) Geeks, social imaginaries, and recursive publics. *Cultural Anthropology* 20(2): 185–214.

Kelty CM (2013) From participation to power. In: Delwiche A and Jacobs J (eds), *The Participatory Cultures Handbook*, Abingdon: Routledge, pp. 22–31.

Kennedy H, Moss G, Birchall C, et al. (2014) *Digital data analysis, public engagement and the social life of methods: Final Report*. Leeds: Communities & Cultures Network. Available from: http://www.communitiesandculture.org/files/2013/02/DDAfinal-report.pdf (accessed 10 September 2015).

Kildall S and Stern N (2011) Wikipedia Art: Citation as Performative Act. In: Lovink G and Tkacz N (eds), *Critical Point of View: A Wikipedia Reader*, INC Reader, Amsterdam: Institute of Network Cultures, pp. 161–190.

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

Kitzinger J (2000) Media templates: patterns of association and the (re)construction of meaning over time. *Media, Culture & Society* 22(1): 61–84.

Kullenberg C and Palmaas K (2009) *Tarde's contagiontology: from ant hills to panspectric surveillance technologies*. Eurozine. Available from: http://www.eurozine.com/pdf/2009-03-09-kullenberg-en.pdf (accessed 7 January 2014).

Laclau E and Mouffe C (2001) *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. London: Verso.

Langlois G, Elmer G, McKelvey F, et al. (2009) Networked publics: the double articulation of code and politics on facebook. *Canadian Journal of Communication* 34(3): 415–434.

Laniado D, Tasso R, Volkovich Y, et al. (2011) When the wikipedians talk: network and tree structure of wikipedia discussion pages. In: *Proceedings of 5th Annual Conference on Weblogs and Social Media*, Menlo Park, CA: AAAI.

Lash S, Szerszynski B and Wynne B (1996) Introduction: ecology, realism and the social sciences. In: Lash S, Szerszynski B, and Wynne B (eds), *Risk, Environment and Modernity: Towards a New Ecology*, London: SAGE, pp. 1–25.

Latour B (1986) Visualization and cognition. *Knowledge and Society* 6: 1–40.

Latour B (1987) *Science in Action: How to Follow Scientists and Engineers Through Society*. Cambridge, MA: Harvard Univ Press.

Latour B (1988) *The Pasteurization of France*. Cambridge, MA: Harvard University Press.

Latour B (1991) Technology is society made durable. In: Law J (ed.), *A Sociology of Monsters: Essays on Power, Technology and Domination*, London: Routledge, pp. 103–131.

Latour B (1992) Where are the missing masses? The sociology of a few mundane artefacts. In: Bijker W and Law J (eds), *Shaping Technology/Building Society: Studies in Sociotechnical Change*, Cambridge, MA: MIT Press, pp. 225–258.

Latour B (1993) *We Have Never Been Modern*. Cambridge, MA: Harvard University Press.

Latour B (1999) *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Latour B (2004) Why has critique run out of steam? From matters of fact to matters of concern. *Critical inquiry* 30(2): 225–248.

Latour B (2005) *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Latour B (2010) Tarde's idea of quantification. In: Candea M (ed.), *The Social After Gabriel Tarde: Debates and Assessments*, Abingdon: Routledge, pp. 145–162.

Latour B and Bastide F (1986) Writing science—fact and fiction: the analysis of the process of reality construction through the application of socio-semiotic methods to scientific texts. In: Callon M, Law J, and Rip A (eds), *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, London: Springer, pp. 51–66.

Latour B and Lépinay VA (2009) *The Science of Passionate Interests: An Introduction to Gabriel Tarde's Economic Anthropology*. Chicago: Prickly Paradigm Press.

Latour B and Weibel P (eds) (2005) *Making Things Public: Atmospheres of Democracy*. Cambridge, MA: MIT Press.

Latour B and Woolgar S (1979) *Laboratory Life: The Construction of Scientific Facts*. London: SAGE.

Latour B, Mauguin P and Teil G (1992) A note on socio-technical graphs. *Social Studies of Science* 22(1): 33–57.

Latour B, Jensen P, Venturini T, et al. (2012) The whole is always smaller than its parts: a digital test of gabriel tarde's monads. *British Journal of Sociology* 63(4): 590–615.

Law J (1986) The heterogenaity of texts. In: Callon M, Law J, and Rip A (eds), *Mapping the Dynamics of Science and Technology*, London: Springer.

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

Law J and Hassard J (1999) *Actor Network Theory and After*. Oxford, UK: Blackwell Publishing.

Law J and Ruppert E (2013) The social life of methods: devices. *Journal of Cultural Economy* 6(3): 229–240.

Law J, Ruppert E and Savage M (2011) *The double social life of method*. Working Paper, Milton Keynes: CRESC, Open University.

Lazer D, Pentland AS, Adamic L, et al. (2009) Life in the network: the coming age of computational social science. *Science* 323(5915): 721.

Lewenstein BV (1995a) From fax to facts: communication in the cold fusion saga. *Social Studies of Science* 25(3): 403–436.

Lewenstein BV (1995b) Science and the media. In: Jasanoff S, Markle GE, Peterson JC, et al. (eds), *Handbook of Science and Technology Studies*, Thousand Oaks: Sage, pp. 343–60.

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

Lezaun J and Soneryd L (2007) Consulting citizens: technologies of elicitation and the mobility of publics. *Public Understanding of Science* 16(3): 279–297.

Lievrouw LA (1990) Communication and the social representation of scientific knowledge. *Critical Studies in Media Communication* 7(1): 1–10.

Lievrouw LA (2011) *Alternative and Activist New Media*. Cambridge: Polity.

Lievrouw LA (2014) Materiality and media in communication technology studies: an unfinished project. In: Gillespie T, Boczkowski PJ, and Foot KA (eds), *Media Technologies: Essays on Communication, Materiality, and Society*, Cambridge, MA: MIT Press, pp. 21–52.

Livingstone S (2007) On the material and the symbolic: Silverstone's double articulation of research traditions in new media studies. *New Media & Society* 9(1): 16–24.

Livingstone S and Lievrouw LA (2009) *New Media*. London: SAGE.

Luke TW (1987) Chernobyl: the packaging of transnational ecological disaster. *Critical Studies in Media Communication* 4(4): 351–375.

Lynch M (1988) The externalized retina: selection and mathematization in the visual documentation of objects in the life sciences. *Human Studies* 11(2): 201–234.

Manovich L (2012) 'Trending: the promises and the challenges of big social data.'. In: Gold MK (ed.), *Debates in the Digital Humanities*, Minneapolis, MN: University Of Minnesota Press.

Manovich L, Douglass J and Huber W (2011) Understanding scanlation: how to read one million fan-translated manga pages. *Image & Narrative* 12(1): 206–228.

Marcus GE (1995) Ethnography in/of the world system: the emergence of multi-sited ethnography. *Annual Review of Anthropology* 24: 95–117.

Markham AN (2005) The methods, politics, and ethics of representation in online ethnography. In: Denzin NK and Lincoln YS (eds), *The Sage Handbook of Qualitative Research*, Thousand Oaks, CA: SAGE.

Marrero-Guillamón I (2013) Actor-network theory, Gabriel Tarde and the study of an urban social movement: the case of Can Ricart, Barcelona. *Qualitative sociology* 36(4): 403–421.

Marres N (2002) May the true victim of defacement stand up! On reading the network configurations of scandal on the web. In: Latour B and Weibel P (eds), *Iconoclash: Beyond the Image in Wars in Science, Religion and Art*, Cambridge, MA: MIT Press, pp. 486–489.

Marres N (2005a) Issues spark a public into being. A key but often forgotten point of the Lippmann-Dewey debate. In: Latour B and Weibel P (eds), *Making Things Public: Atmospheres of Democracy*, Cambridge, MA: MIT Press, pp. 208–217.

Marres N (2005b) No issue, no public: democratic deficits after the displacement of politics. Thesis, Amsterdam: University von Amsterdam. Available from: http://dare.uva.nl/document/17061 (accessed 6 February 2015).

Marres N (2006) Net-work is format work: issue networks and the sites of civil society politics. In: Dean J, Asherton J, and Lovink G (eds), *Reformatting Politics: Information Technology and Global Civil Society*, London: Routledge, pp. 3–17.

Marres N (2012a) *Material Participation: Technology, the Environment and Everyday Publics*. London: Palgrave Macmillan.

Marres N (2012b) The redistribution of methods: on intervention in digital social research, broadly conceived. *The Sociological Review* 60(S1): 139–165.

Marres N (2015) Why map issues? On controversy analysis as a digital method. *Science, Technology & Human Values*: 655–686.

Marres N and Gerlitz C (2015) Interface methods: renegotiating relations between digital social research, sts and sociology. *Sociological Review*.

Marres N and Lezaun J (2011) Materials and devices of the public: an introduction. *Economy and Society* 40(4): 489–509.

Marres N and Moats D (2015) Mapping controversies with social media: the case for symmetry. *Social Media + Society* 1(1).

Marres N and Rogers R (2000) Depluralising the web and repluralising public debate: the case of the gm food debate on the web. In: *Preferred Placement*, Maastrict: Jan Van Eyck Editions, pp. 113–135.

Marres N and Rogers R (2005) Recipe for tracing the fate of issues and their publics on the web. In: Latour B and Weibel P (eds), *Making Things Public: Atmospheres of Democracy*, Cambridge, MA: MIT Press.

Marres N and Weltevrede E (2013) Scraping the social? Issues in real-time social research. *Journal of Cultural Economy* 6(3): 313–335.

McCarthy E and Kelty C (2010) Responsibility and nanotechnology. *Social Studies of Science* 40(3): 405–432.

McCarthy JD and Zald MN (1977) Resource mobilization and social movements: a partial theory. *American Journal of Sociology*: 1212–1241.

Melucci A (1996) *Challenging Codes: Collective Action in the Information Age*. Cambride, UK: Cambridge University Press.

Mendoza M, Poblete B and Castillo C (2010) Twitter Under Crisis: Can we trust what we RT? In: *Proceedings of the First Workshop on Social Media Analytics*, New York: ACM, pp. 71–79.

Meraz S and Papacharissi Z (2013) Networked gatekeeping and networked framing on #Egypt. *The International Journal of Press/Politics* 18(2): 138–166.

Mercea D (2013) Probing the implications of facebook use for the organizational form of social movement organizations. *Information, Communication & Society* 16(8): 1306–1327.

Merton R (1957) Priorities in scientific discovery: a chapter in the sociology of science. *American Sociological Review* 22(6): 635–659.

Merton RK (1973) *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago press.

Michael M (2000) These boots are made for walking...: mundane technology, the body and human-environment relations. *Body & Society* 6(3-4): 107–126.

Miller D, Kitzinger J and Beharrell P (1998) *The Circuit of Mass Communication: Media Strategies, Representation and Audience Reception in the AIDS Crisis*. London: Sage Publications Ltd.

Miller V (2008) New media, networking and phatic culture. *Convergence: The International Journal of Research into New Media Technologies* 14(4): 387–400.

Mol A (1999) Ontological politics. A word and some questions. *The Sociological Review* 47(S1): 74–89.

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

Molotch H and Lester M (1974) News as Purposive Behavior: On the Strategic Use of Routine Events, Accidents, and Scandals. *American Sociological Review* 39(1): 101–112.

Morita A, Blok A and Kimura S (2013) Environmental Infrastructures of Emergeny: The Formation of a Civic Radiation Monitoring Map during the Fukushima Disaster. In: Hindmarsh R (ed.), *Nuclear Disaster at Fukushima Daiichi: Social, Political and Environmental Issues*, London: Routledge, pp. 78–95.

Morozov E (2012) *The Net Delusion: The Dark Side of Internet Freedom*. London: PublicAffairs.

Murthy D (2008) Digital ethnography an examination of the use of new technologies for social research. *Sociology* 42(5): 837–855.

Murthy D (2013) *Twitter: Social Communication in the Twitter Age*. Cambride, UK: Polity.

Murthy D and Longwell SA (2013) Twitter and disasters. *Information, Communication & Society* 16(6): 837–855.

Nelkin D (1971) *Nuclear Power and Its Critics: The Cayuga Lake Controversy*. Ithaca, NY: Cornell University Press.

Nelkin D (1974) The role of experts in a nuclear siting controversy. *Bulletin of the Atomic Scientists* 30(9): 29–36.

Nelkin D (1981) Some social and political dimensions of nuclear power: examples from Three Mile Island. *The American Political Science Review*: 132–142.

Niederer S and van Dijck J (2010) Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system. *New Media & Society* 12(8): 1368–1387.

Nowotny H and Hirsch H (1980) The consequences of dissent: sociological reflections on the controversy of the low dose effects. *Research Policy* 9(3): 278–294.

Olson M (1965) *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

O'Reilly T (2005) What is web 2.0, design patterns and business models for the next generation of software. *O'Reilly.com*. Available from: http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html (accessed 14 September 2015).

Osborne T and Rose N (1999) Do the social sciences create phenomena? The example of public opinion research. *The British journal of sociology* 50(3): 367–396.

Oswell D (2002) *Television, Childhood and the Home: A History of the Making of the Child Television Audience in Britain*. Oxford: Oxford University Press.

Otway H, Haastrup P, Cannell W, et al. (1988) Risk communication in Europe after Chernobyl: a media analysis of seven countries. *Organization & Environment* 2(1): 3–15.

Papacharissi Z (2002) The virtual sphere: the Internet as a public sphere. *New Media & Society* 4(1): 9–27.

Park HW and Thelwall M (2005) The network approach to web hyperlink research and its utility for science communication. In: Hine C (ed.), *Virtual Methods: Issues in Social Research on the Internet*, London: Berg, pp. 171–182.

Peoples C (2014) New nuclear, new security? Framing security in the policy case for new nuclear power in the United Kingdom. *Security Dialogue* 45(2): 156–173.

Perrow C (1984) *Normal Accidents: Living with High Risk Systems*. New York: Basic Books.

Philo G (2007) News content studies, media group methods and discourse analysis: a comparison of approaches. Devereux E (ed.), *Media Studies: Key Issues and Debates*: 101–133.

Pinch T (1994) Cold fusion and the sociology of scientific knowledge. *Technical Communication Quarterly* 3(1): 85–100.

Pinch T and Leuenberger C (2006) Studying scientific controversy from the STS perspective. In: *Proceedings of EASTS Conference 'Science Controversy and Democracy',* Taiwan: National Taiwan University. Available from: http://ionesco.sciences-po.fr/com/moodledata/3/Pinch_Leuenberger_Controversies.pdf (accessed 7 August 2012).

Plantin J-C (2011) *'The map is the debate': radiation webmapping and public involvement during the fukushima issue*. SSRN Scholarly Paper, Rochester, NY: Social Science Research Network.

Plesner U (2009) An actor-network perspective on changing work practices communication technologies as actants in newswork. *Journalism* 10(5): 604–626.

Price V, Tewksbury D and Powers E (1997) Switching Trains of Thought The Impact of News Frames on Readers' Cognitive Responses. *Communication research* 24(5): 481–506.

Procter R, Vis F, Voss A, et al. (2011) How riot rumours spread on Twitter. *the Guardian*. Available from: http://www.theguardian.com/uk/interactive/2011/dec/07/london-riots-twitter (accessed 5 December 2013).

Procter R, Vis F and Voss A (2013) Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology* 16(3): 197–214.

Rabeharisoa V (2003) The struggle against neuromuscular diseases in France and the emergence of the 'partnership model' of patient organisation. *Social Science & Medicine* 57(11): 2127–2136.

Rabeharisoa V, Moreira T and Akrich M (2014) Evidence-based activism: patients', users' and activists' groups in knowledge society. *BioSocieties* 9(2): 111–128.

Reagle Jr JM (2010) *Good Faith Collaboration: The Culture of Wikipedia*. Cambridge, MA: The MIT Press.

Reese SD and Ballinger J (2001) The roots of a sociology of news: Remembering Mr. Gates and social control in the newsroom. *Journalism & Mass Communication Quarterly* 78(4): 641–658.

Rheingold H (1993) *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge, MA: The MIT Press.

Rieder B (2012) The refraction chamber: Twitter as sphere and network. *First Monday* 17(11-5).

Rieder B (2013) Studying Facebook via data extraction: the Netvizz application. In: *Proceedings of the 5th Annual ACM Web Science Conference*, New York: ACM, pp. 346–355.

Rodríguez-Giralt I (2011) Social movements as actor-networks: prospects for a symmetrical approach to doñana's environmentalist protests. *Convergencia. Revista de Ciencias Sociales* 18(56): 13–35.

Rogers R (ed.) (2000) *Preferred Placement: Knowledge Politics on the Web*. 1st ed. Jan Van Eyck Editions, the Netherlands.

Rogers R (2004) *Information Politics on the Web*. Cambridge, MA: The MIT Press.

Rogers R (2009) *The end of the virtual: digital methods*. Amsterdam: Amsterdam University Press.

Rogers R (2012) Mapping and the politics of web space. *Theory, Culture & Society* 29(4-5): 193–219.

Rogers R (2013a) Debanalizing twitter: the transformation of an object of study. In: *Proceedings of the 5th Annual ACM Web Science Conference*, New York: ACM, pp. 356–365.

Rogers R (2013b) *Digital Methods*. Cambridge, MA: MIT Press.

Rogers R, Weltevrede E, Borra E, et al. (2013) National Web Studies. In: Hartley J, Burgess J, and Bruns A (eds), *A Companion to New Media Dynamics*, London: John Wiley & Sons, pp. 142–166.

Sack W (2000) Discourse diagrams: interface design for very large-scale conversations. In: *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, IEEE, p. 10–pp.

Sampson TD (2012) *Virality: Contagion Theory in the Age of Networks*. Minneapolis, MN: University of Minnesota Press.

Sanger L (2005) The early history of Nupedia and Wikipedia: a memoir. In: DiBona C, Cooper D, and Stone M (eds), *Open Sources*, Sebastopol, CA: O'Reilly Media, pp. 307–338.

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(5): 885–899.

Savage M, Ruppert E and Law J (2010) *Digital devices: nine theses*. CRESC Working Paper Series, CRESC University of Manchester, Open University.

Schäfer MS (2012) Taking stock: a meta-analysis of studies on the media's coverage of science. *Public Understanding of Science* 21(6): 650–663.

Schlecker M and Hirsch E (2001) Incomplete knowledge: ethnography and the crisis of context in studies of media, science and technology. *History of the Human Sciences* 14(1): 69–87.

Schmitz Weiss A and Domingo D (2010) Innovation processes in online newsrooms as actor-networks and communities of practice. *New Media & Society* 12(7): 1156–1171.

Shirky C (2009) *Here Comes Everybody: The Power of Organizing Without Organizations*. London: Penguin Group.

Siles I (2011) From online filter to web format: articulating materiality and meaning in the early history of blogs. *Social Studies of Science* 41(5): 737–758.

Silverstone R (1985) *Framing Science: The Making of a BBC Documentary*. London: British Film Institute London.

Silverstone R (1994) *Television and Everyday Life*. Psychology Press.

Silverstone R and Hirsch E (1992) *Consuming Technologies: Media and Information in Domestic Spaces*. Hove, UK: Psychology Press.

Simon B (2001) Public science: media configuration and closure in the cold fusion controversy. *Public Understanding of Science* 10(4): 383–402.

Sismondo S (2008) Science and technology studies and an engaged program. In: Hackett EJ, Amsterdamska O, Lynch M, et al. (eds), *The Handbook of Science and Technology Studies*, pp. 13–32.

Slater D (2014) *New Media, Development and Globalization: Making Connections in the Global South*. London: John Wiley & Sons.

Slater D, Keiko N and Kindstrand L (2012) Social Media in Disaster Japan. In: Kingston J (ed.), *Natural Disaster and Nuclear Crisis in Japan: Response and Recovery after Japan's 3/11*, London: Nissan Institute/Routledge Japanese Studies.

Sloan L, Morgan J, Housley W, et al. (2013) Knowing the tweeters: deriving sociologically relevant demographics from twitter. *Sociological Research Online* 18(3): 7.

Sood BR, Stockdale G and Rogers EM (1987) How the news media operate in natural disasters. *Journal of Communication* 37(3): 27–41.

Star SL (1990) Power, technology and the phenomenology of conventions: on being allergic to onions. *The Sociological Review* 38(S1): 26–56.

Star SL (1999) The ethnography of infrastructure. *American Behavioral Scientist* 43(3): 377–391.

Stengers I (2010) *Cosmopolitics I*. Minneapolis, MN: University Of Minnesota Press.

Sumi R, Yasseri T, Rung A, et al. (2011) Edit wars in Wikipedia. In: *Privacy, Security, Risk and Trust (PASSAT), 2011 IEEE Third International Conference*, Boston, MA. Available from: http://arxiv.org/abs/1107.3689 (accessed 2 July 2015).

Tapscott D (2006) *Wikinomics: How Mass Collaboration Changes Everything*. New York: Portfolio.

Thorpe C and Gregory J (2010) Producing the post-Fordist public: The political economy of public engagement with science. *Science as Culture* 19(3): 273–301.

Thorsen E (2008) Journalistic objectivity redefined? Wikinews and the neutral point of view. *New Media & Society* 10(6): 935–954.

Thorson K (2014) Facing an uncertain reception: young citizens and political interaction on Facebook. *Information, Communication & Society* 17(2): 203–216.

Tkacz N (2010) Wikipedia and the Politics of Mass Collaboration. *Journal of Media and Communication*.

Tkacz N (2012) The Truth of Wikipedia. *Digithum* 14. Available from: http://www.uoc.edu/ojs/index.php/digithum/article/download/n14-aibar-fuster/n14-dossier-en#page=35 (accessed 11 July 2012).

Tkacz N (2014) *Wikipedia and the Politics of Openness*. Chicago: University of Chicago Press.

Toscano A (2007) Powers of pacification: state and empire in Gabriel Tarde. *Economy and Society* 36(4): 597–613.

Touraine A (1983) *Anti-nuclear Protest: The opposition to nuclear energy in France*. Cambridge University Press.

Tuchman G (1978) *Making News: A Study in the Construction of Reality*. New York: Free Press.

Tufekci Z (2013) *Big data: pitfalls, methods and concepts for an emergent field*. SSRN Scholarly Paper, Rochester, NY: Social Science Research Network.

Turner F (2010) *From Counterculture to Cyberculture: Stewart Brand, the Whole Earth Network, and the Rise of Digital Utopianism*. Chicago: University Of Chicago Press.

Turner S (2003) *Liberal Democracy 3.0: Civil Society in an Age of Experts*. London: Sage Publications Limited.

Uprichard E (2013) Focus: big data, little questions? *Discover Society*. Available from: http://www.discoversociety.org/2013/10/01/focus-big-data-little-questions/ (accessed 30 September 2014).

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

Van Dijk TA (1993) Principles of critical discourse analysis. *Discourse & Society* 4(2): 249–283.

Van Loon J (2011) How to be mediatized? An invitation to metaphysics in defense of actor network theory. In: *International Communication Association's 2011 Virtual Confererence*, Raleigh, NC: North Carolina State University.

Venturini T (2010a) Building on faults: how to represent controversies with digital methods. *Public Understanding of Science* 21(7): 796–812.

Venturini T (2010b) Diving in magma: how to explore controversies with actor-network theory. *Public Understanding of Science* 19(3): 258–273.

Venturini T and Guido D (2012) Once upon a text: an ant tale in text analysis. *Sociologica* (3): 0–0.

Venturini T and Latour B (2010) The social fabric: digital traces and quali-quantitative methods. In: *Proceedings of Future En Seine*, CAP Digital, pp. 30–15.

Venturini T, Laffite NB, Cointet J-P, et al. (2014) Three maps and three misunderstandings: A digital mapping of climate diplomacy. *Big Data & Society* 1(2).

Venturini T, Ricci D, Mauri M, et al. (2015) Designing Controversies and Their Publics. *Design Issues* 31(3): 74–87.

Verran H (2012) Number. In: Lury C and Wakeford N (eds), *Inventive Methods: The Happening of the Social*, London: Routledge, pp. 110–124.

Viégas FB, Wattenberg M and Dave K (2004) Studying cooperation and conflict between authors with history flow visualizations. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Montreal Canada: ACM, pp. 575–582.

Vieweg S, Hughes AL, Starbird K, et al. (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, New York, NY, USA: ACM, pp. 1079–1088.

Vis F (2013a) A critical reflection on big data: considering apis, researchers and tools as data makers. *First Monday* 18(10).

Vis F (2013b) Twitter as a reporting tool for breaking news. *Digital Journalism* 1(1): 27–47.

Wajcman J and Jones PK (2012) Border communication: media sociology and STS. *Media, Culture & Society* 34(6): 673–690.

Weart SR (2012) *The Rise of Nuclear Fear*. Cambridge, MA: Harvard University Press.

Welsh I (2000) *Mobilising modernity: The nuclear moment*. London: Routledge.

White M (1950) The gatekeeper. *Journalism Quarterly* 27: 383–90.

Wikipedia Editors (2013a) Wikipedia: neutral point of view. *Wikipedia, the free encyclopedia*. Available from: http://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=531272988 (accessed 13 January 2013).

Wikipedia Editors (2013b) Wikipedia: verifiability. *Wikipedia, the free encyclopedia*. Available from: http://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&oldid=532783572 (accessed 13 January 2013).

Wilkie A, Michael M and Plummer-Fernandez M (2015) Speculative method and Twitter: bots, energy and three conceptual characters. *The Sociological Review* 63(1): 79–101.

Wilkins L (1993) Between facts and values: print media coverage of the greenhouse effect, 1987-1990. *Public Understanding of Science* 2(1): 71–84.

Woolgar S (ed.) (2002) *Virtual Society*. Oxford: Oxford University Press.

Wynne B (1992) Misunderstood misunderstanding: Social identities and public uptake of science. *Public understanding of science* 1(3): 281–304.

Wynne B (1996) May the sheep safely graze? A reflexive view of the expert-lay divide. In: Lash S, Szerszynski B, and Wynne B (eds), *Risk, Environment and Modernity: Towards a New Ecology*, London: Sage, pp. 44–83.

Wynne B (2005) Risk as globalizing 'democratic'discourse? Framing subjects and citizens. *Science and citizens*: 66–82.

Wynne B (2008) Elephants in the rooms where publics encounter 'science'?: A response to Darrin Durant,'Accounting for expertise: Wynne and the autonomy of the lay public'. *Public Understanding of Science* 17(1): 21–33.

Wynne B (2011) *Rationality and ritual*. Abingdon: Earthscan.

Yamamura E (2012) Effect of free media on views regarding nuclear energy after the Fukushima accident. *Kyklos* 65(1): 132–141.

Yaneva A (2013) *Mapping controversies in architecture*. Farnham, Surrey: Ashgate Publishing, Ltd.

Yasseri T, Sumi R, Rung A, et al. (2012) Dynamics of Conflicts in Wikipedia. *PLoS ONE* 7(6). Available from: http://dx.doi.org/10.1371/journal.pone.0038869 (accessed 23 July 2012).

Ziman J (1991) Public understanding of science. *Science, Technology, & Human Values* 16(1): 99–105.

Zimmer M (2010) 'But the data is already public': On the ethics of research in Facebook. *Ethics and Information Technology* 12(4): 313–325.

TECHNICAL APPENDIX

While the rationale and inputs behind individual visualisations have been discussed in individual chapters, this appendix provides detailed settings and procedures so that the different visualisations can be reproduced. Every tool used is freely available with the exception of Microsoft Excel (https://www.office.com/) and Outwit Hub (https://www.outwit.com/products/hub/).

01. LIKE NETWORK

This visualisation was produced with the tool Netvizz (https://apps.facebook.com/netvizz/), developed by Bernhard Rieder. I first created a new Facebook identity (which identified myself as a researcher at Goldsmiths) and 'liked' the following pages related to the topic of nuclear power, mainly in the UK.

Boycott EDF,
CNDUK,
JANUK,
South West Against Nuclear,
Stop Hinkley,
EDF Energy,
Pandora's Promise

These were determined partially by searching 'nuclear power' through the Facebook search – which, based on my location is geographically biased toward the UK – and specific organisations which came up in field notes, flyers and other websites. I also purposely added EDF Energy and *Pandora's Promise* as representatives of the nuclear industry and pro-nuclear environmental position respectively, in order to see how these positions would connect or not with anti-nuclear and environmental groups.

For each of these pages, I used the Netvizz interface to obtain a 'like' network in .gdf format with a depth of 2 – a list of pages which are liked by the starting page and a list of pages liked by those pages and the connections between them.

In network analysis tool Gephi (http://gephi.github.io/) I used the 'append graph' function to join the separate like networks together – nodes with the same identity are

merged, pooling their connections. I set the minimum degree threshold at 2 to include only those pages 'liked' or 'liked by' 2 pages. Note that this is a simpler method than co-link below as it does not distinguish the directionality of links as a marker of authority. This allowed me to keep the starting points in the picture and allowed me to explore variability in the uses of liking. In larger like networks it might be necessary to reduce by 'in degree' to retained only pages repeatedly linked *to*.

The graph was spatialised using Force Atlas 2 (settings: default but with 'lin long mode' and 'prevent overlap' checked). I then used the 'modularity' script in the statistics panel (default settings) to identify 'communities' of nodes and used these grouping to assign colour codings in the partition panel. Nodes were sized by degree, simply the number of edges.

02. ISSUE CRAWLER

For this issue network I used the Issue Crawler interface (https://www.issuecrawler.net) I first obtained the following list of starting points through a virtual ethnographic tracing of links, offline ethnography and Google searches for 'nuclear power, uk'. In this list I tried to get a mix of anti-nuclear websites which were largely talking to each other, nuclear industry websites and government and non-government organisations.

Starting Points:

http://cnduk.org
http://januk.org
http://kicknuclear.org
http://llrc.org
http://mariannewildart.wordpress.com
http://mrwsold.org.uk
http://noend.org.uk/index.htm
http://nuclearlakes.moonfruit.com
http://stophinkley.org
http://stopnuclearpoweruk.net
http://www.british-energy.com
http://www.edfenergy.com

http://www.environment-agency.gov.uk
http://www.greenpeace.org/international/en/campaigns/nuclear
http://www.horizonnuclearpower.com
http://www.hse.gov.uk/nuclear
http://www.hse.gov.uk/nuclear/ocns
http://www.iaea.org
http://www.niauk.org
http://www.nnl.co.uk
http://www.no2nuclearpower.org.uk
http://www.sepa.org.uk
http://www.uk-atomic-energy.org.uk
http://www.world-nuclear.org
https://www.gov.uk/government/organisations/department-of-energy-climate-change
https://www.gov.uk/government/policy-teams/82

I ran the crawl on the 19th of May 2013 (completed on the 20th) with the following settings:

Co-Link analysis by [page]
Iterations [2]
Crawl Depth [3]
Privileged starting points [off]

'Co-link' first scrapes the above pages for all outlinks then reduces the list to those outlinks which are linked to by at least two of the starting points, as opposed to snowball which retains pages receiving at least one link from the seeds and inter-actor which merely shows the relationships between the starting points. Co-link 'by page' was selected because often the relevant material was on a subdomain or specific pages rather than the domain or homepage generally. 'Iterations' describes the number of times this process is repeated. Crawl depth refers to how far outside of the specific starting point page within the page the crawler is allowed to go. 'Privilege starting points' was set 'off' in order to entertain the possibility that pages other than the starting points would be more central to the network.

The resulting network was exported to Gephi. It was spatialised using the circular layout, ordered and coloured, to mirror classic Issue Crawler studies, by type (.org, .gov .com .edu .net etc). Nodes were sized by 'authority' – defined by Issue Crawler as the number of inward links from the network.

03. EDITS VERSUS SIZE

This graph was produced using the DMI's Wikipedia Edits Scraper and IP Localizer Tool: (wiki.digitalmethods.net/Dmi/ToolWikipediaEditsScraperAndIPLocalizer) which downloads the entire edit history page of a given Wikipedia article, including the unique version id, timestamp, user, total size of file, user name, user page, comments accompanying edit, IP address of user. To this file I added a column called 'edit number' which assigned numbers to edits from earliest to latest.

With this data I produced two different graphs using Excel's standard chart interface, which I then overlaid. The first charted number of edits over time in which the x-axis was time and the y-axis represented the cumulative number of edits at a particular time. The second chart showed the total size over time where time was on the x-axis and size (in bytes) was on the y-axis. Please note that while the number of edits only goes up, the cumulative size of the article can either go up or down – one graph is cumulative and the other a total snapshot. Also note that while the two x-axes correspond exactly, the two y-axes on the left and the right were adjusted to fix on the same graph – so they are relative to each other. The overall fit of the two lines demonstrates that most of the activity in the first week led to the expansion of the article, which largely did not contract over time.

04. WIKIPEDIA REFERENCES

First 4 Days

For this visualisation I first started with the same Wikipedia Edits Scraper .csv as above for both the page 'Fukushima Daiichi Nuclear Disaster and 'Fukushima Daiichi Nuclear Plant'

1. Sampling the edits

First, due to the size restrictions I will discuss in a moment, I needed to sample the edits: 1 in every 10 for the first few days. I created a column with the formula **"=IF(RIGHT(A2)="1",1,0)"** to place a 1 next to every 10th edit – every time the edit count (A2) ends in 1 (1, 11, 21, 31… etc.) I then filtered the file to include only every ten edits (with a one next to them). For both graphs I cut off the file after the first 4 days.

I published this .csv file as a webpage in Excel making sure to expand the columns so that none of the content was cut off when rendered in html.

2. Scraping versions for references

I opened this .html as a local file in the program Outwit Hub which is used for scraping data from websites. I selected the column that contained the list of links to individual versions of the article corresponding to the edit number.

e.g. –
http://en.wikipedia.org/w/index.php?title=Fukushima_Daiichi_nuclear_disaster&oldid=418588421

For each of these links I ran a scraper with Outwit Hub which cycled through the code to obtain citation number and corresponding hyperlinks based on their location in the source code. I used the settings

Table Tab> Autoexplore Pages> Fast scraper>Use

Scraper:
Apply if page URL contains http://en.wikipedia.org/w/index.php?

	Name	**Marker Before**	**Marker After**
true	ID	<li Id="cite_note-	"><<
true	Link	class="external text" href="	">

This was based on the assumption that recurring textual elements (underlined) will surround the two bits of data I needed (in bold). For example:

<li id="cite_note-WNN-6">^ ^a ^b Massive earthquake hits Japan World Nuclear News, March 11, 2011 2148h GMT (update 8)

Note that this scraper does not take into account multiple uses of the same exact reference. Many links with have the following format 1^abcdefg to signify different references. This could be incorporated into further analyses but I am focusing here on the decision to include the reference or not, which is different from the amount of mileage the editors get out of the same reference.

I specifically included the citation ID to tell me when a reference in the page *did not* contain a link. The resulting file was exported as another .csv which contained a row for every reference in every page scraped, including reference number, hyperlink, unique edit URL, version, date and time.

3. Preparing the scraped references

To this file I added several columns:
A) The column "domain" to extract the host domain from the specific URL of the reference using the following formula:

> =IF(ISERROR(FIND("http://",C2)),MID(C2,FIND("http://",C2)+4,FIND("/",C2,9)-FIND("http://",C2)-4),MID(C2,FIND("//",C2)+2,FIND("/",C2,9)-FIND("//",C2)-2))
>
> Where C2 = Original URL. So 'http://www.bbc.co.uk/news/uk-politics-33604287' --→ 'www.bbc.co.uk'

References without scrapable URLs were returned as '#VALUE!', which I renamed as 'NO LINK' to draw attention to either the lack of link or failure of the simple scraper to obtain it.

B) I also added a column to group the references by edit number (identifying the version of the article from which they were extracted).

=IF(C3=C2,D2,D2+1)

Where C = Edit URL and D=edit number

C) I also created a value for time expressed as a number. I first converted the column Version Time and Version Date retained from the original Wikipedia Scraper output to a number using **Format>Cells>Number**, allowing 2 decimal places, which converts the dates to Unix time, calculated as number of seconds elapsed since 1 January 1970 UTC. I added the columns together and then subtracted the number of the earliest edit, in this case 39153.38 to give a more manageable measure of time relative to the first edit.

D) Finally I added a column for 'Totals', in which simply put the value 1 next to each reference so that multiple instances of a domain (bbc.co.uk) in a particular version would be added together in the next step.

4. Visualising with RAW

In visualisation programme RAW by Density Design (http://raw.densitydesign.org/) I uploaded the modified csv and selected the 'Stream Graph' visualization which, is essentially an area graph, and assigned the following columns to the inputs. For the field 'Group', I selected the column 'Domain', for 'Date' I entered 'Relative Unix' Time and for 'Size' I entered 'Totals' (the column containing a 1 for each instance of a URL). I chose 'Zero' as the most legible 'Offset' for the graph and visualised it at 2000 x 1500.

At each given time slice (in this case the time stamp of the edit) the total number of instances of each domain are represented as bars of colour with connecting lines. The vertical order is based on the order of the first time slice with new domains being added on top. The total height of the graph is thus absolute.

The same procedure was used for the graph of the total year, except that I used the following formula to obtain every 100 edits – placing a one next to every edit ending in 11 (11, 111, 211 …etc.)
"=IF(RIGHT(A2)="11",1,0)"

There was an error which only manifested itself in the year long span which was that when the user page could not be found by the Wikipedia Edit Scraper (if they had been banned from Wikipedia for example) and there was no value in this field then the scraper failed to return references for the edits before and after. This only impacted 5 edits over the course of the year (2611 , 4511, 5311, and 5711) which were manually removed. So the graph merely skips these edits appearing to jump ahead slightly. However the Raw Stream graph positions the lines in real time so this only affects the granularity not the rhythm of the data. This can hopefully be corrected in future versions.

05. BI-PARTITE NETWORKS

*See earlier in the appendix for description of like network, also featured in this chapter.

1. Email List Scraping

A) To scrape data from the RiseUp Email List I used a chrome extension called Scraper (search 'scraper' in Chrome Web Store). I first selected the tab on the email list interface to sort posts 'chronologically' rather than 'threaded' which became easier to scrape. On the first page (which represents all the messages for a month) starting from the date of the Fukushima Disaster, I right clicked on one of the posts and selected "scrape similar" starting the Scraper application.

The scraper identified the location of the posts in the source code (in XPath markup) as:

```
//div[4]/div/div/ul/li/ul/li
```

Starting from this location I further specified the following items within the "li" tag also expressed in X Path

./b/a	Post Title
./em	User
./b/a/@name	Message Number
./b/a/@href	Message URL

../../strong Message Date

This gave me a .csv file for the current page and I manually repeated this process for each month of the year following the Fukushima disaster March 2011 – April 2012 and I then cut and paste the .csvs into one file.

I first removed the ubiquitous text '[KickNuke]' from the posts:

=IF(ISNUMBER(SEARCH("[KickNuke]",A2)), REPLACE(A2, H2, 11, ""), "?")
Where A is the column with the original post title

Then I separately determined if the post was a reply or not by asking if the post started with "re:":

=IF(ISNUMBER(SEARCH("re:",I2,1)), SEARCH("re:",I2,1), "notreply")
Where I is the Column with the post – [kicknuke]

If it was a reply I deleted the "re: " at the beginning of posts in order to make the text of originals and replies comparable.

=IF(ISNUMBER(J2), REPLACE(I2, J2, 4, ""), I2)
Where J is the answer "noreply" or "1" for a reply

2. User-Post Bi-Partite

Using this CSV I used Table 2 Net to create a bi-partite network (http://tools.medialab.sciences-po.fr/table2net/index.php). I uploaded the .csv from the previous step and selected bi-partite as the type of network (two types of nodes). The first type of nodes I defined as Text, that is the title of the post and for the second type of nodes I defined as User.

I exported the file to Gephi, spatialised it using Force Atlas 2 (again, lin long mode, prevent overlap checked). Nodes were coloured by type (user or post) and sized by degree (number of connections).

3. Word-Post Bi-Partite

For the co-word network I first needed to convert the text of the posts into separate text files. I imported the 'Post Title', 'User' and 'Unique Identifier' columns into a separate excel file and created a column 'FileName' which combined the 'Unique ID-Author'. I then ran the following Macro (courtesy of Lilith Whittles).

```
Sub Export_Text3()


Const My_Path1 = "Users:davidmoats:Dropbox:Goldsmiths"
  Dim iCol As Integer
  Dim lRow As Long
  Dim i As Integer
  Dim File_Num As Long
  Dim SaveDest As String

  On Error Resume Next

    If Trim(Dir(My_Path1, vbDirectory)) = "" Then
      MkDir My_Path1
    Else
      Kill My_Path1 & "*.txt"
    End If

  On Error GoTo 0

  For i = 1 To Range("N").Value
  File_Num = FreeFile
  With ActiveSheet
    Open Trim(Range("FileNames")(i, 1).Value) & ".txt" For Output As
#File_Num

      Print #File_Num, Range("Text")(i, 1).Value

    Close #File_Num
  End With
  Next i
End Sub
```

For each row of the table (each post in the email list) this macro created a unique .txt file with the title FileName containing the post text.

This resulting folder full of .txt files was then imported into ANTA (http://jiminy.medialab.sciences-po.fr/anta_dev/index/login) for analysis. In the ANTA beta test interface I uploaded all of the .txt files being careful to note any files

which failed to upload. Some of the .txt files failed to load if they contained non UTF-8 characters so these characters needed to be manually removed on occasion. Some of these were Japanese characters, which is potentially significant given the strong presence of Japanese language groups in the UK. This was helped by saving the file in Libre Office .odt format which handles non-UTF8 characters. (https://www.libreoffice.org/) before bringing it back into Excel.

I clicked the analysis button to the left to begin the term extraction. This extraction is based on a stop list and the Alchemy database which recognises known proper names (including alternate spellings of these names). This is a necessary step but not an unproblematic one because it assumes that different spellings or lesser known entities will be less consequential. Nonetheless the process does pick up unknown words which appear multiple times in the data set. The Include Entities Tab at the top is normally used for filtering terms but I preferred to leave all of the terms in. The output of ANTA is a bi-partite graph containing the original text files (documents) and terms appearing in them, connected by edges.

I then exported the file to Gephi where I removed terms that only appeared in one document (minimum degree 2), spatialised the graph using Force Atlas 2 (lin long mode, prevent overlap), coloured the nodes by type (documents, terms) and sized by degree.

Facebook – Bi-Partite Networks.

The procedure for creating bi-partite networks of Facebook pages is similar but with a few key differences.

1) User-Post Bi-partite Graph

To obtain a network of users and posts in Facebook I used Netvizz, selecting 'Page Data' and entering the desired date range (2012-3 For Stop New Nuclear and Stop Hinkley). I then clicked the button for 'posts by page and users' which includes posts by the page admin as well as posts created by other users. This supplied files containing stats about each post, a .tsv file containing the text of each post and all the comments and a bi-partite graph containing users and posts that they create, comment on, share or like.

The network file was imported into Gephi and spatialised with Force Atlas 2 (lin long mode, prevent overlap), coloured the nodes by type (users, posts by page and posts by user) and sized by degree.

2) Word-Post Bi-partite Graph
The word post graph starts with the .tsv file created in the last step containing the text and comments of the post

In this case the comments needed to be removed leaving only a file containing the original post text, the user who posted it, and the unique identifier of the post (e.g. -228971750452013_573044446044740) which represents page id followed by post id, something which will come in handy in the next step. In the excel macro, this number becomes the file name of the text files and is retained as the node name when ANTA converts it to a network.

The network file was imported into Gephi and spatialized with Force Atlas 2 (lin long mode, prevent overlap), coloured the nodes by type (documents, terms) and sized by degree.

06. TRI-PARTITE NETWORKS

1. Formatting the User-Post Graph

The tri-partite networks are simply a combination of the above two bi-partite graphs, joined at the posts. The trick is to add in time. Fortunately the bi-partite User-Post graph supplied by Netvizz already contains a column called Unix Time which can be used for this purpose. Unix gives a numeric value for dates – the number of seconds elapsed since 1 Jan 1970. To this I added a column called 'Axis' in which I assigned 1 to all user nodes and 2 to all word nodes.

To visualise the time dimension, in Gephi I used the 'Spatial Layout' plugin (search within the interface for plugins). In the Rank panel, clicking the Spatial Layout logo, I selected the y-axis and assigned the column Unix Time between 0 – 3000 and for the x-axis by Axis between 0 and 500. This placed all of the posts in one column and all of the words in a separate column and sorts all of the posts by their unix time. Since space in

Gephi is entirely relative it these numbers are in relation to each other. I sized the nodes by degree and coloured them by type of node (user, post by user, post by page).

In order to better position the user nodes, which of course do not have a time dimension, I first selected the posts, now forming a vertical line down the middle and from the right click menue chose 'settle'. This locks the nodes into place. I then ran Force Atlas 2 which allowed the user nodes to flow freely around the fixed posts – gradually drifting towards posts they commented on more.

Although this arrangement is somewhat arbitrary and more of a visual aide, the problem is that users who engage with a large number of posts, rather than a few posts intensely, will be drawn to the middle of the graph. To combat this I returned to the rank panel and this time assigned the x-axis to Out-degree or the number of times a node connects to something else. This meant that the users who engage with more posts are pushed away from the graph. The posts remain in a line because their out degree value is by definition 0.

2. Formatting the Word-Post Graph

First in Gephi with the Word-Post graph open I clicked the tab for Data Table and clicked Nodes and Export as .csv. This produced a list of all the nodes in the Gephi File as a csv table. I first created an extra column which extracted the unique identifier of the posts from the ANTA created filename: 40-47207552617990.txt →
281732915171_47207552617990
47207552617990

Then I sorted the nodes by type and unique identifier and manually pasted in the extra columns from the Netvizz tab file (engagement, unix time, user etc). I also added a column called Axis and gave every post a value of 1 and every term a value of 2. Then in Gephi I imported the modified file, ensuring that I checked all the boxes of the new columns I had added and selected 'string' for text fields and 'float' or 'integer' for numbers. Sometimes these properties were not recognized on import but this could be fixed by making a duplicate of the column and re-specifying integer or float.

Again in the Rank Panel I assigned y-axis to Unix Time between 0 – 3000. I again sized nodes by degree and coloured by type. I also applied the out-degree ranking as with the previous map but, reversed the order 500 and 0

3. Combining the two graphs

Finally the two graphs were brought into Adobe Illustrator and manually overlaid at the posts which were, in theory, in exactly the same place.

07. SOCIO-TECHNICAL GRAPHS

1. The data set

The staff at the Digital Methods Initiative were kind enough to give me a query bin for my project which from 9 March 2013 onwards collected tweets containing the terms: Fukushima, nuclear, nuke, Hinkley, EDF. The terms nuclear and nuke were general enough that they contained many other controversies over nuclear weapons, nuclear power plant proposals in other parts of the (English speaking) world and even jokes involving someone 'going nuclear'. Hinkley and EDF were far more issue-specific, though there were still Tweets about Hinkley the place generally and similar acronyms to EDF.

As well as offering statistics on top users, top hashtags and top URLs, and of course the facility to download collections of Tweets based on sub-queries and time frames, the interface also gives volume over time with which I was able to determine various spikes in activity, keeping in mind that some of these spikes pertained to non topic specific discussions, as noted earlier. These needed to be investigated manually as discussed in the chapter

2. Obtaining the links

As mentioned in the chapter, URLs in Twitter are now automatically truncated using a t.co link. This builds on the common practice of users using link shortening services like tiny.url. The effect is that links can be truncated, then truncated again when they

are retweeted and truncated again by the retweeter. Add to this the fact that links are not specifically recognised by the API as separate elements from the text of the tweet.

At my suggestion, Erik Borra wrote a script which extracted the URLs from the tweet and then unpacked them until it arrived at a non-truncated original URL, placed into a separate column. If there were two URLs in a single tweet they were separated by a comma. In the interface, an option was added to search by URL or part of URL.

3. Frequencies

For each of the three URLs discussed, I first obtained every instance of them in a .csv file using the query function giving the full date range of the data set to make sure I got them all. I was careful to use only the main text of the link as an identifier because some links would have the addition /m/ or /mobile to identify the version of the same article viewable on a phone. I used the service shared count to make sure I got most of the links, even if they did not contain the keywords necessary to be included in my data set.

To produce a frequencies graph, I simply extracted the time stamps of all the tweets for each URL, adding a column for Number (the number of the Tweets in order) and the name of the URL. I brought this file into RAW and selected 'Scatterplot'. For the x-axis I assigned Timestamp, for the y-axis I assigned Number and for Colour I assigned URL. This graph shows the total number of shares of a URL over time, particularly the relative rate at which they advance and in what rhythm.

4. Colour Coding

The first step in producing the coloured strips was to first remove the URL from the text of the tweet – with truncations the URL would vary considerably and I was specifically interested in what else would change. This could be accomplished by the following formulas to identify the character position of 'http:' and the next available blank space after and thus the length of the URL

```
=FIND("http",AE2,1)
=FIND(" ",AE2, AG2)
=AG2-AF2
```

Where AE is the column with the text of the tweet and AG is the character position of "http" and AF is the end of the URL. The URL was then removed with the formula:

=REPLACE(AE2,AF2,AG2," ")

I also had to manually remove URLs where there were more than one in a single tweet.

Next I needed to assign Number to unique tweets.First, the following formula runs through the tweets to determine if a tweet has been used before or not:

=COUNTIF($AI$2:AI2,AI2)

Where AI is the column containing the text of the tweet without the URL.This counts the number of times a tweet has appeared previously in the series. If it has never appeared the value will be 1. In the next formula every unique tweet, maked with a 1 is given a new unique number starting with one. If it has appeared before (anything other than one, then it looks up through the list to find the number of the first instance of the tweet it matches.

=IF(AJ574=1,MAX($AK$1:AK573)+1,INDEX($AK$2:AK574,MATCH(AI574,$AI$2:AI574,0),1)

Where AJ is the number of times a tweet appears, AK is the column assigning the unique identifier (where this formula appears) and AI is the text of the tweet. The result is that if the Tweet is unique then the numbers advance but if the tweet has been used before it locates the first instance of the tweet and assigns that number.

To apply the colours to this I used the conditional formatting menu which assigns a colour along a three-colour spectrum according to the unique number obtained in the last step. Please note that to make the links comparable then needed to be loaded first into a single column so that the colours and numbers would correspond exactly, rather than relatively within each URL.

08. URL SEQUENCER

The URL Sequencer tool is currently being developed with the help of Erik Borra at the DMI. The current incarnation builds on the approach used above but rather than simply identifying original or repeated Tweets, distinguishes typologies of tweets.

This works by stripping away some of the incidental formatting to indentify the base text. This includes removing 'RT @_____ ' 'via @_____' '@_____' or '@_____', lower case the full tweet; replace the following with a space: {:, cc:,",',...,-}; remove trailing spaces; replace all sequences of spaces by a single space.

Given a column of all the tweets in time order, the remaining base text is then assigned unique colours when it is repeated. Perfectly unique tweets remain on a white background without formatting for easy identification.

The tool then assigns a unique column to each tweet typology from left to right in the order in which they first appear. Although this does not give a sense of time, it very quickly gives the researchers a sense of the order of major events and the extent to which a particular URL is relatively homogenous or heterogeneous in its output.

Although this is still under prototype the other function of the tool will be to identify variations within the base tweet – if a hashtag or @ mention is added to the text for example. This would require setting a threshold of how many characters or words would constitute a modification, but not a brand new tweet. Within this threshold a script similar to the diff function (https://en.wikipedia.org/wiki/Help:Diff) used in Wikipedia would highlight in green words that are added and highlight in red words that are removed.

This tool will be detailed in a forthcoming joint paper with Erik Borra.