London South Bank University

From the SelectedWorks of Philip Hammond

December 15, 2015

From Computer-Assisted to Data-Driven: Journalism and Big Data

Philip Hammond

Available at: http://works.bepress.com/phil_hammond/12/
From Computer-Assisted to Data-Driven: Journalism and Big Data

Philip Hammond
London South Bank University

Abstract
Despite claims of continuity, contemporary data journalism is quite different from the earlier tradition of computer-assisted reporting. Though it echoes earlier claims about being scientific and democratic, these qualities are understood as resulting from better data access rather than as being something achieved by the journalist. In the context of Big Data in particular, human subjectivity tends to be downgraded in importance, even understood as getting in the way if it means hubristically theorising about causation rather than working with correlation and allowing the data to speak. Increasing ‘datafication’ is not what is driving changes in the profession, however. Rather, the impact of Big Data tends to be understood in ways that are consonant with pre-existing expectations, which are shaped by the broader contemporary post-humanist political context. The same is true in academic analysis, where actor-network theory seems to be emerging as the dominant paradigm for understanding data journalism, but in largely uncritical ways.

Keywords: actor-network theory, Big Data, computer-assisted reporting, data journalism, objectivity, post-humanism

THIS IS AN AUTHORS’ DRAFT. THE FINAL VERSION OF THIS ARTICLE WAS PUBLISHED IN JOURNALISM, ONLINE FIRST: 15 DECEMBER 2015 (DOI: 1464884915620205) HTTP://JOU.SAGEPUB.COM/CONTENT/EARLY/2015/12/15/1464884915620205.ABSTRACT
From Computer-Assisted to Data-Driven: Journalism and Big Data

Philip Hammond
London South Bank University

There is no shortage of terms describing something new going on in contemporary journalism. Many of them – such as data-driven journalism, computational journalism, programmer journalism, algorithmic journalism, robot reporting, automated journalism – are very directly related to the use of data and computers; and other, somewhat broader, characterisations – such as interactive journalism, networked journalism or post-industrial journalism – suggest a similar emphasis. Of course there are distinctions to be made, and some scholars have helpfully sought to clarify the meaning of particular terms in relation to somewhat different practices of using data and computers in journalism (see, for example, Coddington, 2015). Yet it is the fact of so many new expressions being coined to point in a broadly similar direction that seems striking, rather than the more subtle differences between them.

This article argues that while journalism is indeed changing, digital data and computer technology are less central to understanding this than is often assumed. If one wanted to sum up this change in terms of a shift in terminology, then the contrast between journalism which is ‘computer-assisted’ and that which is ‘data-driven’ would be a good indicator: computers and data, rather than the human subject, have come to be understood as the active party in the relationship. Although it is of course true that how we do journalism has changed, the more important change is in how we think about journalism, and that change is not simply driven by new technologies but needs to be understood in broader contextual terms. Recent debates about data journalism, particularly against the background of the rise of Big Data, crystallise these changes. ‘Big Data’ is still a relatively new concept and its implications remain uncertain, but its proponents claim that the datafication of our everyday lives, through our interactions with ‘smart’ devices and sensors which can communicate with each other, will improve our understanding of what is happening in the world and thereby enable us to make better decisions. Importantly, it is claimed that this understanding will come not from human theorising and hypothesis-testing but from correlations revealed by computer processing of data. Rather than digital data technologies simply causing changes in journalism, however, it is more that both Big Data and (data) journalism are increasingly understood in ways that are consonant with broader shifts in how we think about the human subject and his/her ability to know about and act on the world as object. These broader changes are also evident in what seems to be an emerging consensus in the scholarly engagement with data journalism which, rather than offering a critique, finds confirmation of its own post-humanist assumptions.

Continuity and change

Despite the profusion of terminology highlighting new journalism practices, there is also a tendency to emphasise continuity with the past, perhaps partly because of a desire to bolster their legitimacy as ‘proper’ journalism. Simon Rogers, for example, who established the Guardian’s Data Blog in 2009, argues that the newspaper has been doing data journalism since 1821, when its first edition carried a table of data about Manchester Schools (Rogers, 2011). The ‘big difference’, he argues, is that data used to come in the form of ‘very expensive books’, whereas now there are ‘spreadsheets and files formatted for computers’. The change is understood to be basically technical, a matter of old analogue formats being superseded by new digital
technologies, though importantly his comments also highlight the greater accessibility of digital data. The most commonly invoked point of historical comparison, however, is not 19th century newspapers but the ‘computer-assisted reporting’ (CAR) which emerged in the US in the late 1960s and early ‘70s. The relationship between CAR and contemporary data journalism has been widely debated. While not everyone would agree with the National Institute for Computer-Assisted Reporting that the only thing to have changed is the name,1 many commentators have suggested that there is an underlying continuity linking newer innovations with a past tradition (Bounegru, 2012; Lewis and Westlund, 2015). Yet while some characterisations of data journalism can indeed sound very similar to CAR, on closer inspection there is a marked difference between the two. So much so, indeed, that data journalism might almost be understood as the antithesis of CAR.

Although there were earlier precedents, it was Philip Meyer’s 1973 book Precision Journalism which announced and advocated the arrival of data analysis as a significant part of the journalistic repertoire. Effectively a manifesto for CAR, Precision Journalism above all promoted a ‘scientific’ approach: ‘The new precision journalism is scientific journalism… adopting the scientific method, scientific objectivity and scientific ideals’ (Meyer, 2002: 5). In an apparent echo of Meyer, in 2010 Julian Assange described his ambition for Wikileaks as being ‘to set up a new standard: “scientific journalism”’ (in Khatchadourian, 2010). Yet as Lisa Lynch has astutely observed, Meyer and Assange have very different understandings of what constitutes ‘scientific journalism’:

> while Meyer argues that looking at data is intended to confirm that ‘scientific method’ has resulted [in] journalism free of interpretive error…[for] Assange, journalists should not be expected to produce a consistent accounting of the truth: rather, the truth is to be determined by the individual reader through the consultation of source documents.

(Lynch, 2012: 53)

Whereas, for Meyer, better methods of assembling and analysing source materials can enable journalists to achieve a more scientific approach to reporting, Assange argues that, since he makes his source materials available to the reader, he is free to add whatever commentary or interpretation he likes. Although it is sometimes claimed that part of the contemporary appeal of data journalism is that it is seen to offer a ‘renewed promise of objectivity’ (Parasie, 2015: 365), it is better understood, at least in Assange’s version, as part of a ‘post-objective’ journalistic culture, in which, as Lynch (2012: 54) suggests, the profession is more comfortable with a goal of ‘transparency’ rather than ‘objectivity’.

The transparency of much data journalism – laying bare the sources and the data – is often linked to the news media’s traditional watchdog role, in which informing the public sphere and holding power to account are central. At first glance, this again appears to be a point of continuity with CAR, which is closely associated with investigative, ‘civic’ or ‘public service’ journalism. As Meyer explains:

> As originally conceived during the social protest movements of the 1960s, precision journalism was a way to expand the tool kit of the reporter to make topics that were previously inaccessible, or only crudely accessible, subject to journalistic scrutiny. It was especially useful in giving a hearing to minority and dissident groups that were struggling for representation.

(Meyer, 2002: 235)

Making a similar claim for today’s data journalism, Liliana Bounegru argues that:
By enabling anyone to drill down into data sources...data journalism effectively represents the mass democratisation of resources, tools, techniques and methodologies that were previously used by specialists – whether investigative reporters, social scientists, statisticians, analysts or other experts.

(Bounegru, 2012)

Just as the scientificity of data journalism is held to be given by the audience’s relationship to the source material, so too the democratising role of data journalism is understood not so much as deriving from what journalists themselves do but instead as consequent on the openness and accessibility of data sources.

A case in point is the Guardian’s coverage of the 2009 British MPs’ expenses scandal. Having been left behind, like everyone else, by the Daily Telegraph’s scoop on this story, the Guardian made what the Nieman Journalism Lab described as a ‘spectacular’ recovery, by crowdsourcing its reporting (Andersen, 2009). The paper released the entire dataset of expenses claims on its website and asked its readers to help analyse the documents and identify potential stories for Guardian journalists to investigate further (Rogers, 2009). The invitation to readers to join in with the investigative journalistic process could be understood as an ingenious way to catch up quickly with a rival paper, but also, less cynically, as exemplifying just the sort of democratising impulse that Bounegru sees in much data journalism. Yet this example also clearly highlights a dilemma for professional journalists. On the one hand, journalism that consists of publishing databases then adding comment pieces and/or asking the audience to lead the investigation diminishes both claims to journalistic authority and the need for professional journalists. Yet on the other hand, given that journalistic authority is already in question, given the fuzzy boundaries between professionals and amateurs, and given the declining audience for paid-for journalism in the West, the professionals also feel compelled to be transparent and open, and to acknowledge the audience as co-producers. Hence, in his ‘10 point guide to data journalism’, Rogers finds himself making contradictory claims. He argues that:

There’s now so much data out there in the world that we try to provide the key facts for each story – and finding the right information can be...a lengthy journalistic task...

[We act as the bridge between the data (and those who are pretty much hopeless at explaining it) and the people out there in the real world who want to understand what that story is really about.

(Rogers, 2011)

Yet having established that identifying ‘key facts’ is a substantial ‘journalistic task’ (point three), to be undertaken by specialists who can really explain the story (point five), Rogers then feels compelled to add as point seven that ‘Anyone can do it’. In fact point seven of his 10 point guide ends with an ellipsis, and is taken up in point eight: ‘...but looks can be everything’ – suggesting that the professionals can still add something important in terms of sophisticated data-visualisations and graphics, designed by people who ‘understand the issues involved’. Yet here too, it seems, anyone can do it: as well as publishing multiple datasets for readers to investigate, the Guardian also invites readers to join in with creating infographics, asking them to ‘Please post your visualisations and mash-ups on our Flickr group’.2 If one rationale for data journalism is that we are suffering from a data deluge and need the professional journalist to help us make sense of it all, another is that the data revolution allows journalism to become transparent and open-source, making it
possible for anyone and everyone to have first-hand access to the raw material of the story. In the first scenario, data is opaque and confusing, requiring special expertise to interpret it and turn it into meaningful knowledge; in the second, data is transparent and the only requirement is that we be given access to it. The first argument is more similar to Meyer’s proposition for CAR: that equipping journalists with new knowledge and skills would enhance their expertise and their special ability to explain events for their audience. Yet today’s data journalists, perhaps partly out of concern that the public may be paying less and less attention to professional journalism, often find themselves drawn instead to the idea that new data-driven practices offer a way to engage an audience by making journalism into something that everyone can join in with doing.

Underlying this uncertainty about the relationship between professional journalism and the public is a broader question about the nature of knowledge and interpretation in an era of Big Data. The comparison with CAR is again instructive here. For Meyer, knowledge is clearly understood to be hypothesis-driven, a matter of theorising and interpreting:

You can’t begin to think about a problem without some kind of theoretical framework….The process of hypothesis formulation in scientific method forces the framework to the conscious level where it can be coldly and objectively evaluated.

[W]ithout theory, we have nothing but unordered raw data, and we suffocate in it.

(Meyer, 2002: 9, 13)

For the evangelists of Big Data, in contrast, the point is precisely that it makes theory obsolete, an unnecessary barrier to ‘letting the data speak’. As Wired editor Chris Anderson famously put it:

Out with every theory of human behavior, from linguistics to sociology….With enough data, the numbers speak for themselves….We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot….The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.

(Anderson, 2008)

In Anderson’s telling, theorising is made obsolete by the new availability of massive datasets and vastly increased computing power. What is at stake in the move from human interpretation and hypothesis-testing to machine-processing of data is an epistemological shift, from understandings of the world based on working out theories of causation to knowledge based on correlation. The argument that, as Anderson (2008) puts it, ‘With enough data, the numbers speak for themselves’ implies that the only reason we used to try to work out models, formulate hypotheses and understand causation was that we lacked the affordances of Big Data. Now, it seems, we are fortunate to be able to leave behind the quest to understand the world in terms of cause and effect. This implies very different expectations of journalism and its capacity to explain events.
In a 1980 article partly inspired by CAR, and Meyer’s description of *Precision Journalism* as offering ‘a reporter’s introduction to social science methods’, David Weaver and Maxwell Combs traced the history of the relationship between journalism and social science, seeing then-recent trends as welcome confirmation that the two had grown closer together, but also arguing that the relationship had deeper roots. They cited, for example, Curtis MacDougall’s 1938 textbook, *Interpretative Reporting*, which maintained that: ‘To interpret the news it is necessary to understand it, and understanding…involves recognizing the particular event as one of a series with both a cause and an effect’ (in Weaver and Combs, 1980: 487). What seemed true in 1938 seemed equally valid in 1980: that a more social-scientific approach to journalism was a positive development because it could do more than simply registering the welter of daily events, and could dig down to discover underlying causes, thereby offering a deeper understanding of the social world. Today, in contrast, even the aspiration to understand cause and effect is starting to look hopelessly antique. As Kenneth Cukier and Viktor Mayer-Schoenberger (2013) argue, in an era of Big Data ‘we will need to give up our quest to discover the cause of things, in return for accepting correlations’. Again the driver of this change is understood to be technological innovation: the ‘rise of Big Data’ is said to be profoundly reshaping our understanding of the world, leading us to focus on ‘what’ is happening rather than asking ‘why’ (Cukier and Mayer-Schoenberger, 2013). Yet while there may be, so to speak, a correlation between the datafication of the world and a more widespread sense that theorising society’s causal laws is hubristic and mistaken – a sense that, in Cukier and Mayer-Schoenberger’s (2013) words, ‘many times, when we think we have identified [causation], it is nothing more than a self-congratulatory illusion’ – it does not follow that the first development has led to the second.

‘Human-assisted reporting’

Of course, not all data journalism is related to Big Data, and indeed some of what is labelled as ‘data journalism’ only tangentially or superficially involves using any sort of data, often employing data in ways that are ‘as much decorative as informative’ (Knight, 2013). The aspect of data journalism that is most directly related to Big Data specifically is what is sometimes referred to as ‘automated’, ‘algorithmic’ or ‘robot’ reporting. Although this has been growing for a number of years, it came to widespread attention in March 2014 when a report by ‘Quakebot’ published in the *Los Angeles Times* caught the public eye (Oremus, 2014; Rutkin, 2014). Quakebot, an algorithm written by *Los Angeles Times* journalist Ken Schwencke, turned the data supplied by the United States Geological Survey’s automated earthquake notification service into a brief news report. In under five minutes, the story was checked by Schwencke and uploaded to the newspaper’s website. One of Schwencke’s colleagues described the process as ‘human-assisted reporting’ (in Marshall, 2013).

More common than this in-house operation is the contracting of such services from companies such as Narrative Science, who offer to ‘transform data into stories and insight’, or Automated Insights, who promise to ‘Let your data tell its story’. Narrative Science estimated in 2012 that within fifteen years more than 90% of news would be written by algorithm; a prediction that perhaps seemed outlandish at the time (Levy, 2012). Judged on the sheer volume of reports, however, it is already true. Automated Insights alone produced 300m pieces of content in 2013, a number expected to rise to over a billion the following year (Roose, 2014), and the company claims to have the capacity to produce 2,000 articles per second (Miller, 2015). Just in terms of the number of items of content, this is already ‘more than every other media outlet in the
world combined', and constitutes the 'biggest feat of news production in the history of the world' (Roose, 2014). The rationale behind the production of very high numbers of stories is that they can easily be customised for particular audiences, even for individual readers, in a way that would not be practicable for human journalists. Moreover, while the claim that 'In 20 years, there will be no area in which Narrative Science doesn't write stories' (in Levy, 2012) remains questionable, it is true that both companies have diversified their output, beginning with sports journalism and then moving into business and financial reporting. Narrative Science writes financial reports for **Forbes**, while Automated Insights provides a similar service for the Associated Press news agency.

Of course, despite the startlingly high volume of automated stories, algorithmic reporting still remains only a minor aspect of journalism considered in the round. Yet it does highlight some difficult questions for the profession. Firstly, as Matt Carlson (2015: 426) points out, ‘the shift of news from a collective statement about the relative importance of events meant for wide audiences to personalized accounts…redefines journalism from a collective statement to individualized information’. The proliferation of customised stories for niche audiences implies that even if Western news organisations were to reverse the historic decline in audience interest, that would not necessarily entail any reinvigoration of a common public sphere. Secondly, automated journalism seems to make a more reliable claim to factuality, neutrality, and objectivity than human journalists. One line of argument against the ideal of journalistic objectivity was that it was an unattainable goal precisely because of human fallibility. As Ivor Gaber (2008) has argued, for example, ‘It must surely be self evident that objectivity is, and has always been, a meaningless concept…because all journalists – subject to official confirmation – are human beings’. If some journalism is now not written by human beings, however, then perhaps it can achieve ‘a meticulous commitment to factuality…precisely because it is not human’, offering ‘a new form of knowledge creation exceeding human capabilities’ (Carlson, 2015: 427). From October 2014 Associated Press dropped the initial policy of having each algorithmically-generated story checked by a human journalist before publication (Miller, 2015), and both Automated Insights and Narrative Science claim that automated reporting produces fewer errors and inaccuracies than would be the case with human writers.

Thirdly, if much journalism is so routine and conformist that it can be written by algorithm, and if machines do objectivity better than humans ever could, then what room is there for a human contribution? Some writers have attempted to address this question by defining the value of human journalism in different terms, emphasising qualities such as performance and emotionality. C.W. Anderson, Emily Bell and Clay Shirky, for example, argue that:

> very personal and human activities mark journalism as a form of information performance rather than simply a dissemination of facts….Cultural literacy skills distinguish reporters, editors, designers and other journalists from other systems of data gathering and dissemination.  
>  
> (Anderson, Bell and Shirky, 2012: 28-9)

Similarly, Matt Waite asserts combatively:

> You know what there isn’t an algorithm for? Humanity….Great journalism, I believe, reflects us as human beings: flawed, complicated, emotional. It is precisely the things that can’t be defined in a programming language that makes us human.  
>  
> (in Carlson, 2015: 428)
Yet flawed, complicated and emotional are qualities that are more or less the opposite of what, in the past, we would have said were the characteristics of ‘great journalism’. It is perhaps worth also noting that again this is very different from the precision journalism approach of CAR: Meyer (2002: 3-4) explicitly contrasted the scientific method he advocated with the New Journalism of Tom Wolfe and others who ‘pushe[d] journalism toward art’. Today it perhaps seems appealing to opt, if not for art, then for cultural literacy and emotional sensitivity because computers are felt to be better at things like precision, accuracy and objectivity.

It could be argued, then, that Big Data will potentially have a number of important consequences for journalism. If automated reporting continues to develop and grow, a datafied world could effectively report on itself, with less and less need for professional human journalism. If data capture systems embedded into our everyday interactions are increasingly able to determine our interests and preferences, and link this information to the delivery of customisable news content, we might expect that the ideal of a common public sphere, a shared space of democratic deliberation informed by journalism, would increasingly be replaced by small, only partially overlapping, informational sphericules. And human values might come to seem less important and central, if human subjectivity is understood in terms of partiality and fallibility rather than its world-shaping, history-making potential.

Big Data will not cause these changes, however, mainly because they have already happened, for reasons that have nothing to do with Big Data. It may now seem plausible to argue that algorithms are better at objectivity and humans ought to stick to the personal and emotional, for example, but that is not really a new development: the profession has been retreating from objectivity for years. All references to ‘objectivity’ were deleted from the Society of Professional Journalists’ code of ethics in the mid-1990s (Mindich, 1998: 5), for example; around the same time that BBC war correspondent Martin Bell declared that he was ‘no longer sure what “objective” means’ (Bell, 1998: 18). Bell’s alternative approach – what he called the ‘journalism of attachment’ – influenced a generation of war reporters, who often seemed to find a way to make international conflict and intervention all about them, their feelings and their consciences; while more confessional, emotional and personal styles also became increasingly common in other areas of journalism (Coward, 2013; Hume, 1997; Mayes, 2000). Similarly, if Big Data seems to raise the possibility of the fragmentation of the public sphere, we might recall that in the 1990s Todd Gitlin was already discussing how journalism had come to address, not a unitary public sphere, but ‘separate public sphericules’ (Gitlin, 1998). And if open-source data seems to encourage the view that ‘anyone’ can do (data) journalism, we should remember that the BBC already declared that ‘news coverage is a partnership’ after the 7/7 London bombings in 2005, in response to the importance of eyewitness camera-phone pictures, and in the context of professional journalism’s declining audience and authority (Matheson and Allan, 2009: 101). Even the seemingly futuristic idea of a datafied world reporting on itself can be understood as a Big-Data version of much older ideas about the capitalist market as a ‘site of veridiction’ (Foucault, 2008: 32) – as in Friedrich von Hayek’s ‘view of the market as a kind of gigantic information processor superior to highly limited human knowledge or the meddling of political actors’ (Dean, 2014: 438). Rather than Big Data causing far-reaching changes in journalism, it is more that we interpret new developments as confirming or making inevitable what we already thought about it.
Uncritical approaches

If we were to do today what Meyer tried to do in the early 1970s and ‘introduce reporters to social science methods’, it would look rather different. It would have a lot less to do with precision, and more to do with diffuseness, indeterminacy and, as John Law (2004) puts it, ‘mess’. We would be telling reporters that, in a world of complexity, theory-building based on linear cause-and-effect understandings is reductionist. We would be counselling them to abandon any hubristic anthropocentric assumptions they may have and to instead embrace post-humanism and recognise the agency of objects-as-actants. At least, we would if we were actor-network theorists, new materialists, post-humanists, speculative realists or object-oriented ontologists. Such currently fashionable theoretical approaches do indeed seem to be developing into something like the dominant paradigm for understanding data journalism. The academic journal Digital Journalism published two special issues in 2015, on ‘Theories of Journalism in a Digital Age’ and ‘Journalism in an Era of Big Data’, in each of which at least half the articles adopted the same theoretical perspective: actor-network theory (ANT). The common scholarly adoption of ANT as the theoretical framework to analyse data journalism suggests that it has started to feel commonsensical; the most obviously appropriate approach. Why might that be?

One reason is that the sorts of epistemological claims for Big Data discussed above are a good fit for ANT, which finds a kind of self confirmation in the datafied world. Yet if ANT is indeed emerging as the preferred approach for analysing data journalism, it is not as a critical approach but as something closer to a celebration. Bruno Latour et al. (2012: 600) acknowledge the limitations and inequalities of ‘most present day databases’, but, having dropped this curtsey to critique, are effusive in their account of the possibilities that such databases hold. For some years now, Latour (2007) has argued that online digital data potentially offer a powerful vindication both of ANT and of the sociology of Gabriel Tarde from which it draws inspiration. Much like the argument that we only had to go in for social theorising because we lacked the advantages that Big Data brings, Latour suggests that:

the very idea of individual and of society is simply an artifact of the rudimentary way data [were] accumulated. The sheer multiplication of digital data has rendered collective existence (I don’t use the adjective social anymore) traceable in an entirely different way than before.

(Latour, 2011: 803)

In the past, it was very difficult, if not impossible, to ‘drill down’ into aggregate data to the level of individual people, or to trace the multiple connections between those individuals and the networks within which they were embedded. Hence, ‘the notion of society [was] generated, a special way to grasp collective phenomena…that you find just as well in the tired old cliché that “the whole is superior to the sum of its parts”’ (2011: 803). Latour wants to insist with Tarde that, to the contrary, ‘The whole is always smaller than its parts’ (Latour et al., 2012). The supposedly larger category of ‘the social’, he argues, was just a way of coping with the fact that we could not trace the myriad connections between actors; and the notion of ‘the individual’ was a reductive way to deal with our inability to discern how people were embedded in complex networks. As Latour puts it:
To believe in the existence either of individual or of society is simply a way to say that we have been deprived of information on the individuals we started with; that we have little knowledge about their interactions; that we have lost the precise conduits through which what we call ‘the whole’ actually circulates.

(2011: 805)

Now, however, thanks to ‘new digital tools’ enabling us to navigate a newly-available ‘mass of quali-quantitative data’, we can leave those empty categories behind (2011: 807, 809). Fortuitously, this is also a perfect confirmation of ANT, according to which every actor is a network and vice versa.

For journalism studies, too, there is a clear congruence between an ANT approach and the sorts of developments in automated reporting discussed above: if journalism is dependent on computers and data, and in some cases is no longer something done by humans at all, then ANT’s concern with assemblages of humans, objects and technologies is again a perfect fit. Alex Primo and Gabriela Zago (2015: 45), for example, present the new technologies and practices of algorithmic reporting as the best illustration of their argument for the importance of adopting an ANT perspective on journalism. This is only part of the appeal, however – after all, ANT predates such innovations, and the argument is about how we think of journalism in all its historical forms and all its technologies. It is supposed to be equally applicable to the era of well-chewed pencils and shorthand notebooks, or to that of eighteenth-century coffee houses and hand-operated printing presses, as to today’s world of networked computers and massive electronic datasets. As Primo and Zago (2015: 44) emphasise, the significance of ANT is not that it foregrounds the importance of technology (though it does do that), but rather that it has radical ‘epistemological and methodological consequences’ for our understanding of journalism. Their objective is to challenge what they see as a widespread failing: ‘both the humanities and the sociological tradition lead to the same ontological mistake: the supposition that journalism is a practice restricted to humans’ (Primo and Zago, 2015: 46). They argue that the problem with ‘the prevailing humanistic ontology of journalism’ is that it ‘ignores the active participation of non-humans’, and warn that ‘while non-humans are kept as predicates, not as subjects, the anthropocentric tale of journalism will not be surpassed’ (2015: 49, 46). According to Latour, modernity’s anthropocentric outlook never really worked in the way it claimed. It depended on reductionist, binary divisions between Society and Nature, Subject and Object, to construct an heroic tale of progress and revolution, a ‘vast saga of radical rupture, fatal destiny, irreversible good or bad fortune’ (Latour, 1993: 48). Now that modernity’s systems of scientific and political representation have broken down and its grand narratives no longer sound credible, he argues that we can reinterpret the past and see the hybrids and monsters, the networks and assemblages, the middle-ground that modernity studiously ignored (1993: 47). If an ANT account of journalism focuses on ‘the role played by technological artifacts, understood as full-blown social actors, with transforming roles’, then, the point is not simply to ‘highlight the importance of technology’ but to get us to reassess our view of the human subject as the central, active agent (Primo and Zago, 2015: 39, 44).

As Andrew Calcutt and I have argued elsewhere, modern journalism emerged around the same time as modern philosophical ideas of Subject and Object, as a product of the Age of Enlightenment and of mercantile capitalism (Calcutt and Hammond, 2011). The modern, humanist view of the active subject, investigating and transforming the world as object was always in tension, though, with a view of the subject as determined and constrained by objective conditions. As Karl Marx (1852) put it, ‘Men make their own history, but they do not make it as they please;
they do not make it under self-selected circumstances, but under circumstances existing already, given and transmitted from the past. Journalistic objectivity had a similarly double-edged character: in part about a genuine extension of public knowledge and informing public debate; but also partly about narrowing debate within acceptable parameters and burying active subjectivity beneath the dead weight of facts. In the era of nineteenth-century industrial capitalism, different sections of the press tended to divide between radical journalism supporting the active history-making subject, and an alienated objectivity which packaged propaganda for the status quo as ‘objective journalism’ and served it up to a passive/pacified audience. Twentieth-century experiments such as New Journalism or CAR were both reactions against this stifling tradition of alienated objectivity, seeking better ways to ‘communicate essential truth’ (Meyer, 2002: 4). For both, the subjectivity of the journalist was vital, whether understood in terms of artistic licence or scientific discipline. As we have seen, however, today’s data journalism tends to be rather different: echoing earlier claims about being scientific and democratic, but understanding these as resulting from better data access rather than as being something achieved by the journalist. In the context of Big Data in particular, human subjectivity tends to be downgraded in importance, even understood as just getting in the way if it means hubristically theorising about causation rather than working with correlation and allowing the data to speak. In journalism, as in politics, we are no longer in the era of the active, history-making subject. If it is true that, as suggested above, we tend to interpret the impact of Big Data in ways that are consonant with our pre-existing expectations, then ANT embodies those expectations mainly because it describes the world in terms which render it recognisably post-humanist (Chandler, 2015).

**Conclusion: post-human politics**

For Latour (2014: 17), the ‘crucial political task’ of our times is ‘to distribute agency as far and in as differentiated a way as possible – until…we have thoroughly lost any relation between those two concepts of object and subject that are of no interest any more’. The proposition that these are not just theoretical but political arguments is echoed in Primo and Zago’s (2015: 38-9) presentation of an ANT perspective on journalism as an iconoclastic confrontation with previous assumptions, which they characterise as ‘deterministic’, ‘essentialist’, ‘partial’, ‘biased’ and ‘ideologically driven’. Just as some definitions of journalism seek to restrict it to something done only by professionals, they argue, so too the idea that it is ‘restricted to humans’ needs to be questioned (2015: 46-7). To a post-humanist sensibility, an anthropocentric view of journalism, which casts it as a solely human activity, is seen as unacceptably exclusivist and undemocratic, and their understanding of it as created by ‘hybrid collectives’ of humans and non-humans represents the more progressive approach (2015: 43). The idea that expanding our understanding of journalism to include the contribution of non-human actants is a progressive and democratising move – similar to opening journalism up to a wider range of human actors, but even more thoroughgoing and radical – is absurd from a humanist point of view, but makes perfect sense from an ANT perspective. Latour has long pitched his arguments in similar terms, calling for ‘a democracy extended to things themselves’, or for a ‘parliament of things’, for example (Latour, 1993: 142).

The greater digital connectivity between people and objects which underpins the rise of Big Data may understandably look attractive to writers who see it as a vindication of their theoretical perspective, but the notion that it is also a politically desirable and progressive phenomenon is by no means restricted to academic ANT theorists. While the term ‘Big Data’ has had a sinister ‘Big Brother’ ring to it since Edward
Snowden’s 2013 revelations about the US National Security Agency’s surveillance programmes, the infrastructure that enables increasing datafication – the ‘Internet of Things’ – is frequently still discussed in glowing terms. Nowhere is this more evident than in the corporate world, where business-oriented ‘digital consultancies’ and ‘thought leadership platforms’ extol the political virtues of the Internet of Things. It is claimed that ‘The Internet of Things has the potential to vastly strengthen the democratic values and institutions that together form the bedrock of Western civilisation’, and can ‘make government and governance more participatory than ever’ (Ampofo, 2015), for example. Or that ‘The most amazing promise that the Internet of Everything will bring is an extreme sensation of connection between both humans and things that will ultimately blur the sensation and mindset of separation’ (Fonseca, 2015). If there is a faintly Latourian ring to the idea of blurring the subject/object division in a world in which devices and data are ‘intelligent agents’ (Ampofo, 2015), an even more direct translation of ANT into mainstream business discourse is offered by IBM, who rename the ‘parliament of things’ a ‘device democracy’ – a ‘new and flat democracy’ in which ‘power in the network shifts from the center to the edge’, and devices become ‘equal citizens’ which are ‘empowered’ to act ‘autonomously’ (Brody and Pureswaran, 2015: 8-9).

Advocacy of greater participation and connection may sound vaguely progressive, but as Chris Peters and Tamara Witschge (2015: 31) have pointed out in relation to digital journalism, such ideas should not be taken at face value: ‘While digital tools may promote connection in a very literal sense, these are not the same forms of “public connection” emphasised under the old grand narratives’. As they argue, a ‘participation paradigm’ appears to be emerging in terms of how journalism is ‘viewed by academics, spoken about by news organisations, and engaged with by audiences’, and although some of the terminology sounds the same, we should not assume any continuity or congruence with previous assumptions about journalism and democracy (Peters and Witschge, 2015: 24). Democratic participation no longer necessarily means joining in with the formation of a collective public will through democratic debate and deliberation, since today we are more likely to think of ‘politics’ as something that happens in the personal sphere, where plural publics are understood to be self-constituting through everyday associational interaction (Chandler, 2014: 162). And whereas the modern outlook understood the human subject as able to understand and act to transform the world as object, in today’s non-linear world being ‘connected’ does not imply a social or political connection forged between autonomous subjects, but rather signals the way that we are understood to be embedded in assemblages of both human and non-human actors. It is not obvious that acknowledging non-humans as co-participants and co-creators of the world is of any great benefit to the objects, but it does demote the human subject to just another actant. As Peters and Witschge’s (2015) emphasis on intellectual frameworks suggests, although technological innovations, commercial imperatives and other factors no doubt play an important part in changing how we think about journalism and democracy, there is a larger social and political context. It is the collapse of those ‘grand narratives’ of modernity that has allowed a version of post-humanist perspectives such as ANT to become part of mainstream discussion.  

These contextual changes have important implications for our understanding of journalism and its potential to articulate collective meaning. Nick Couldry argues that:

It is easy to give up on agency in a world where so many of our acts are fed into predictive models that have no interest in meaning.
...a polity based on an impoverished model of the human subject cannot expect much loyalty from, or legitimacy with, those it governs.

(Couldry, 2014: 891, 894)

Given that he is one the rare writers to engage in a genuinely critical way with the implications of Big Data it might seem churlish to quibble with him, but the chain of cause and effect does seem to be the wrong way round in this formulation. It is more that models with no interest in meaning seem useful in a world in which we doubt the agency of the subject, and that an impoverished understanding of the human subject seems an obvious basis for a polity which does not command much loyalty from, or legitimacy with, those it governs. That is to say, with the exhaustion of the modern framework of Left and Right, disconnected elites look to other mechanisms of governance – ‘nudging’ the public to adopt particular ideas or behaviours (Thaler and Sunstein, 2008), or seeking to address voters’ emotional wellbeing (Richards, 2007) spring to mind as examples – which reduce or even bypass the need to persuade the public with political arguments or engage them through reasoned debate. Big Data is now emerging as another such governance technique (Chandler, 2015: 835).

As Couldry (2014: 892) argues, in a world of Big Data there seems to be little or no scope for human ‘interpretative agency’. In the twentieth century, collective meaning-making was dominated by large, top-down institutions of representation and mediation such as national broadcasters or newspapers of record. In the early twenty-first century, these started to give way to new bottom-up forms of participatory self-representation by spontaneous collectives formed via social media. These different configurations spawned what Couldry (2014: 881) calls the ‘myth of the mediated centre’ and the ‘myth of “us”’, respectively: different ways of thinking about and representing our mediated relationship to society. The ‘myth of big data’, in contrast, ‘challenges the very idea that the social is something we can interpret at all’ (2014: 882). As we have seen, this myth can be expressed in claims that human interpretation based on theories of causation is unnecessary, or is even a hindrance to letting data correlations show us what is happening; in versions of data journalism in which audiences access data directly, with little or no need for the mediation of professional journalists; and in academic approaches which understand theoretical constructs such as ‘individual’ and ‘society’ as poor substitutes for the sort of thick description now facilitated by digital data. There is a long tradition of critiquing the first myth and challenging the ways that people were (mis)represented by mass media. The initial utopian hopes that online media offered more authentic possibilities for self-representation by the demos were superseded quite quickly by a more downbeat assessment of how powerful new interests were starting to dominate the World Wide Web, allowing a similar critique to be made of how apparently spontaneous self-expression and sharing was actually shaped by large companies such as Facebook or Google. We are, in other words, practiced at deconstructing regimes of collective representation and the meanings they offer. However, we are less sure how to deal with a post-representational regime, in which collective meaning is bypassed in favour of the promise of access to the reality revealed by datafied everyday interactions. Critics who go on relentlessly attacking the modern humanist subject today are apparently oblivious to the fact that they are jabbing away at a corpse. It may be time to consider whether the task of critique might now be better served by a more constructive project, rather than further deconstruction.
Notes

1 According to the National Institute for Computer-Assisted Reporting, ‘The term “computer-assisted reporting” was used widely in the last two decades to describe what many now call “data journalism”’. See: www.ire.org/nicar/about/.

2 Until June 2013 the Guardian made datasets available at www.theguardian.com/news/datablog/interactive/2013/jan/14/all-our-datasets-index. The Flickr group is at www.flickr.com/groups/1115946@N24/.

3 Slogans from the companies’ websites at www.narrativescience.com and http://automatedinsights.com, respectively. While Automated Insights’ slogan encapsulates the idea of ‘letting the data speak’, Narrative Science’s promise implies that data need to be worked on and transformed into a narrative before they can offer insight. The difference is not so great, however. While some data journalists might understand story-telling as a uniquely human contribution, for Narrative Science the transformation is not performed by human journalists but by algorithms.

4 We Have Never Been Modern (Latour, 1993) is the key ‘crossover’ work: the book that allowed ANT to move from the specialist niche of science and technology studies to the grander terrain of social and political theory as confidence in the modern project collapsed at the end of the Cold War.

References


Marx, Karl (1852) *The Eighteenth Brumaire of Louis Bonaparte* [www.marxists.org/archive/marx/works/1852/18th-brumaire/ch01.htm].


