Nick Couldry, Aristea Fotopoulou and Luke Dickens

Real social analytics: a contribution towards a phenomenology of a digital world

REAL SOCIAL ANALYTICS:

A CONTRIBUTION TOWARDS A PHENOMENOLOGY OF A DIGITAL WORLD

8556 words

NICK COULDRY (LSE), ARISTEA FOTOPOULOU (SUSSEX UNIVERSITY)

LUKE DICKENS (OPEN UNIVERSITY)

ABSTRACT

This article argues against the assumption that agency and reflexivity disappear in an age of 'algorithmic power' (Lash 2007). Following the suggestions of Beer (2009), it proposes that, far from disappearing, new forms of agency and reflexivity around the embedding in everyday practice of not only algorithms but also analytics more broadly are emerging as social actors continue to pursue their social ends but mediated through digital interfaces: this is the consequence of many social actors now needing their digital presence, regardless of whether they wants this, to be measured and counted. The article proposes 'social analytics' as a new topic for sociology: the sociological study of social actors' uses of analytics not for the sake of measurement itself (or to make profit from measurement) but in order to fulfil better their social ends through an enhancement of their digital presence. The article places social analytics in the context of earlier debates about categorization, algorithmic power, and self-presentation online, and describes a case study with a UK community organization which

developed the social analytics approach in detail. The article concludes with reflections on the implications of this approach for further sociological fieldwork in a digital world.

KEYWORDS

Social analytics; algorithms; algorithmic power; phenomenology; categorization; agency; reflexivity

REAL SOCIAL ANALYTICS:

A CONTRIBUTION TOWARDS A PHENOMENOLOGY OF A DIGITAL WORLD

In the digital world social actors are used to having their presence in online space categorised and measured. Web analytics and information sorting based on big data (including data derived from social media sources) is now fundamental to how capitalism works, while 'visibility' is, arguably, a key dimension of the contemporary social terrain (Brighenti 2007; Heinich 2012). Adobe's term 'social analytics' refers to uses of such data, gathered particularly via social media interfaces, for commercial value. We appropriate the term here to capture something of independent sociological interest. The sociology of knowledge has always been concerned with 'the analysis of the social construction of reality' (Berger and Luckmann 1967: 15), but any sociology of knowledge today must address the background role of digital infrastructure (algorithms, analytics, architectures, platforms) in shaping what counts as 'reality'. This requires – hence the term *social* analytics – detailed accounts of how social actors construct their own online presence and develop a reflexive relationship to the infrastructures of digital presence. Such detailed accounts constitute what we call 'real social analytics', and contribute to a phenomenology of the digital world.

A social analytics approach – more precisely, a sociological treatment of how analytics get used by a range of social actors in order to meet their social ends – aims to capture how particular actors reflect upon, and adjust, their online presence and the actions that feed into it, through the use of 'analytics'. We use 'analytics' here in an extended sense to mean in this context: 'basic analytics' (the automated measurement and counting installed within the operation of digital platforms, and associated websites, apps and tools); adjustments made by actors themselves so as to incorporate basic analytics into their daily practice; and thirdly, the

architecture (the underlying organization of data flows) that allows digital platforms, and the measuring processes associated with them, to become (through a 'front end' design) embedded in the actions of those who interact with such platforms. Platforms that count and sort online data, such as Google and Facebook, work automatically via algorithms, allowing users only limited degrees of manual adjustment (van Dijck 2013). Other adjustments around those operations may take direct digital form (a website redesign aimed to generate more or better data of a certain sort) *or* organizational form (adjustments in an organization's offline resources, for example of information gathering and storing). The underlying data's relationship to an organization's online presence may be more or less direct: direct if the data is literally *about* that organization's online presence (numbers of unique users, types of interaction with online content); or indirect if about other aspects of performance but generated through an organization's online presence. By 'social actors' we mean actors (individuals or organizations, small or large) with social ends over and above making profit directly from the production, use or sale of basic analytics. Our definition of 'social analytics' therefore encompasses, but goes beyond, the everyday technical use of the term 'analytics' (as 'the measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing Web usage'),[1] to capture all those aspects of the practical operation and reflexive adjustment of the computing-based infrastructures on which digital presence now relies, and that relate to pursuing, reflecting upon and adjusting for, the measurement and evaluation of a social actor's operations online.

While social analytics (so defined) can be pursued from many angles (organizational sociology, economic sociology, political sociology), we argue for its particular interest within a contemporary sociology of knowledge. Social actors today face many pressures to use analytics (in our broader sense) in order to shape how they exist for and are known by others;

their presence and intentionality as social actors is intertwined with the operations of analytic measures and underlying computing architecture. While the mutual intertwining of human and material agency is hardly a new insight (Pickering 1995: 15-20), it acquires special bite when analytics' operations are frequently opaque to non-experts and hard for them to control, even if they do see them at work; such tension is increased for those social actors who are orientated to goals that are distinctively social, such as community organizations, charities, and civil society actors**.** A social analytics approach makes a distinctively qualitative contribution to the expansion of sociological methods in a digital age (Rogers 2010). Sociology must address the changing role of data in everyday life (Halford, Pope and Weal 2012) and resulting new forms of the reproduction of inequality (Halford and Savage 2010). Part of that response will involve new forms of quantitative analysis on datasets large and small (Rogers 2013; Marres and Weltevrede 2013), but equally important is the study of how social actors *themselves* deal with the increasing embedding of quantification, measurement and calculation in their everyday lives and practices. Our starting-point is not so much a 'web epistemology' (Rogers 2010) as a renewed attention to everyday engagements with the epistemological questions raised by the digital infrastructures through which social actors are, increasingly, coming to know themselves.

The concept of social analytics, as a contribution to the sociology of knowledge, emerged from a project of multi-stream action research focused on digital infrastructures' role in facilitating narrative exchange.[2] We were as interested there in the constraints experienced by social actors, as in the opportunities. We reflect here therefore on the obstacles encountered in our fieldwork derived from a gap between our partner's aspirations to have a more effective digital presence, and its resources to implement this, but we also look beyond them towards a wider programme for future research.

Our empirical entry-point should be contextualized in earlier debates about digital sociology and digital research, for example Kling's (2007) vision of 'social informatics' for the algorithmic age. From another angle, our work contributes to debates about algorithms' consequences for the building-blocks of social theory. While the deep embedding of algorithms in social knowledge (Halavais 2008, Lash 2007, Burrows 2009, Gillespie 2014) is not in doubt, we are sceptical of claims that algorithmic technology at work within everyday phenomena creates 'a collapse of ontology and epistemology' (Lash 2006: 581), installing a power-laden regime of 'facticity' (Lash 2007: 56) in which 'there is no time, nor space . . . for reflection' (Lash 2002: 18). Rather, we side with David Beer's claim (in response to Lash) that sociology must also 'focus . . . on those who engage with the software in their everyday lives' (Beer 2009: 999). This allows us to register a key site of *tension* in a digital age when social actors are struggling precisely to make effective use of digital infrastructures in particular contexts shaped by social goals. Evidence of such struggles is now emerging in multiple areas: medical sociology, management studies, sociology of politics.[3]

This is not accidental, but the result of how deeply digital infrastructures now impact on a fundamental question for phenomenology: how 'we derive our sense of self from the image of our self that others reflect back to us in interaction' (Crossley 2001: 143, summarising Cooley 1902). In the digital age, that image of self (individual or collective) may come from other human beings with whom we interact, but also from the whole material infrastructure of calculation, data-processing, and data-presentation through which social actors now appear to others. Studying the consequences of this requires social analytics. In the next section, we unpack in more detail the conceptual background and some conceptual issues that underlie this approach.

Doing phenomenology under digital conditions

Social analytics as a project emerges in a world where many social actors have no choice but to interact with each other, mediated in part by representations of themselves online. Online presence has become a necessity for such actors as they pursue social ends. The resulting proliferation of online presences makes it very difficult to be *distinctively* visible online, to have an *effective* online presence (Hindman 2009). Effective presence requires presence-*to-search* (Introna and Nissenbaum: 2000: 170). Presence-to-search is however only just the beginning, since such presence usually comes with evaluations, whether self-sustained or externally imposed. Such evaluations depend on analytic measures and an underlying infrastructure of data exchange. The encounter with analytics in our broad sense acquires potentially existential importance (Turkle 2011).

Interacting with classifications

The implications of our encounters with computer-generated data within public and private life have increasingly been noticed (Lash 2007; Beer 2007; Burrows 2009; Introna 2011; Van Dijck 2014). Search engines use algorithms to sort vast amounts of data; so too do the fast-growing data-mining industries linked to marketing and advertising (Turow 2011) in a major shift in the production of economic value (Simmel 1990).

Search is a systematic and automatic operationalization of data's requirement for a hermeneutic context: algorithms fill in the context that are needed for the raw data of digital presence to generate information about the existence of a social actor. Social actors have little

choice but to respond to the consequences of such automatic operations, because they directly shape their conditions of presence-to-others. A gap opens up between the automated exploitation of accumulated data about us and our capacity to convert that data back into 'information' usable for our own ends (Kallinikos 2009: 235). As Kallinikos (2009) points out, 'data' only ever become 'information' when they are interpreted in a context that is defined relative to the interests of particular actors. There need be no correspondence between our social goals, and the aggregated data outputs that underlie how our online presence is evaluated: this conflict is exacerbated by the opaqueness to non-specialists of the algorithms underlying most search (Halavais 2009) and other operations of analytics (van Dijck 2013) and computer infrastructure generally.

These transformations – key starting-points for the project of social analytics - are important from at least two broader perspectives. From a historical perspective, the interconnectedness and infinite expandability of the internet - and the massive investment in its searchability - have transformed the information that social actors give off about themselves and their actions. In the digital age, unless actors make considerable efforts *to prevent* an action leaving traces online, those traces will become *de facto* searchable by anyone, anywhere, and for any purpose.[4] This shift, which we are currently living through, may prove as fundamental as the move in Europe's later Middle Ages from oral societies to societies where everyday life came to depend 'for [its] daily business on written record' (Clanchy 1993: 334). It also has more recent precedents: the explosion of information to track work processes that fuelled the 19th century origins of modern management (Chandler 1977; Yates 1989).

From a sociological perspective, this shift is only an expansion, deepening and institutional reinforcement of an abiding feature of social life: categorization (Durkheim and Mauss 1970)

or ‘sorting things out’ (Bowker and Star 1999). Categorization is fundamental to all forms of organization, including social organization. Without it, effective (non-random) interaction with the world would be impossible. In the social realm, categorization has a distinctive feature, because, as Ian Hacking has pointed out, classifications of human objects are ‘interactive’ in a way that classifications of non-human objects are not:

> Ways of classifying human beings interact with the human beings who are classified. There are all sorts of reasons for this. People think of themselves as of a kind, perhaps, or reject the classification. All our acts are under descriptions, and the actions that are open to us depend, in a purely formal way, on the descriptions available to us. Moreover classifications do not exist only in the empty space of language but in institutions, practices, material interactions with things and other people . . . Interactions do not just happen. They happen within matrices, which include many obvious social elements and many obvious materials ones. . . people are aware of what is said about them, thought about them, done to them. They think about and conceptualize themselves. Inanimate things are, by definition, not aware of themselves in the same way. . . . The classifications of the social sciences are interactive. The classifications and concepts of the natural sciences are not. (Hacking 1999: 31-32)

Hacking’s insight has complex implications in the digital age. First, advertisers, marketers and the fast-growing industry that sells data to them are using highly sophisticated tools to classify people in ways of which those people are *not*, at least directly, aware (Turow 2007; Van Dijck 2013); there is controversy on whether citizens should have a right to be *made aware* of such classifications (Turow 2011). Conversely, social actors can now *choose to measure themselves* (and their digital presence) in specific ways and interact with the

classifications that result, so monitoring and modulating their digital presence and its effects. Here Hacking's general insight into the interactive nature of human classifications helps us see a particular, but neglected, aspect of agency in the digital age. Social actors who know their online presence is evaluated, may - indeed are likely to - *want* that presence, in all its modalities, to match their overriding goals, requiring action to achieve this matching.

Social analytics becomes therefore an important site of struggle by social actors to retain control over their conditions of existence, through reflexive adjustment to the parameters of their online categorization and measurement. Social analytics is the sociological study of this struggle. The question of how social actors (mainly individuals) manage their online presence in social networking sites (SNSs) such as Facebook and Twitter has been variously approached; Gerlitz and Helmond (2014) examine the passage from a 'linking' economy (connections between websites) to a 'like' economy, where users gain social currency from the public articulation of connections on social networking sites. In Second Life, the cybernetic social lives of residents are led by their avatars (Boellstorff 2008), whereas in online dating sites users often choose to state that they are younger than they really are (Ellison, Heino, and Gibbs, 2006). Teenagers today take measures to protect their privacy online by making content meaningful only to those they wish to (boyd 2014). Social analytics extends this scholarship on how individuals control their online presence to cover organizations which may have a number of social goals, and complex processes for achieving them. One key dynamic here is evaluation: the generation and application of value through judgement. Contemporary societies are characterised by multiple 'regimes of evaluation' (Boltanski and Thévenot 2006), but paradoxically forms of automated measurement are now also functioning as practical mechanisms for evaluating people and organizations. Social

analytics contributes to wider sociologies of evaluation by exploring how social actors *work through* the consequences of those automated evaluations.

Everyday interactions with the 'back-end' of the digital world

A distinct field of digital sociology is emerging which examines new forms of agency and information politics within the infrastructures of a burgeoning digital realm, particularly their invisible, 'back-end' and automated nature (Marres and Weltevrede 2012; Rogers 2010). Such work follows in the wake of the social studies of technology (STS) tradition, which has always emphasised that technologies are performative, rather than having linear determining effects (MacKenzie and Wajcman 1999; Suchman 2007). Other recent work has encouraged dialogue specifically between the study of code and software and broader sociological and humanities research (Fuller 2005, MacKenzie 2006, Berry 2012). Such work specifically acknowledges the performative nature of code itself: 'code, the material that lies at the core of software, is unstable because it is both expression and action, neither of which are materially or socially stable. In saying something, code also does something, but never exactly what it says' (MacKenzie 2006: 177); 'encoded material enactments translate/extend agency, but never exactly' (Introna 2011: 113).

While acknowledging the importance of such work, we explore here through the notion of social analytics a phenomenological perspective which has been less emphasised within the broader STS tradition. This starts from the premise that so-called 'offline' contexts are an irreducible part of what people do 'online' or in relation to the 'online' (Slater and Miller 2000, noted by Rogers 2010: 242). Qualitative approaches such as ethnography here make an important contribution to understanding people's relations to the 'back-end' of the digital

world: as XXXXb (forthcoming) notes in her mapping of issue- and social networks of London-based feminist organisations, ethnography is especially important for registering how, for example, online link patterns reflect hierarchies and power relations that influence agenda-setting on the ground. More than that, we need research that recognises people's ongoing reflexivity about their conditions of entanglement with digital infrastructures. When, as with 'analytics', actors' mode of appearance-to-others is at stake (a key question for phenomenology), actors may reflect at length on how to change such analytics' operations: in doing so, they may become ever more deeply entangled in digital infrastructures through attempts to influence their outcomes. Such efforts will acquire an additional intensity when other actors seek to impose *their* reading on the outcomes, for example as a way of managing staff (Orlikowski and Scott 2014) or monitoring the 'value' of funded projects (our case study). Such ongoing processes of reflexive engagement with the infrastructures of digital presence are of sociological importance and cannot be adequately dealt with through system-focussed notions of 'apparatus' (Scott and Orlikowski 2013: 78).

In this respect, a phenomenologically sensitive social analytics has much to contribute to the critical study of 'audit culture' (Power 1997), but also to a broader sociology of knowledge. Earlier phenomenology of the social world which aimed to reground a sociology of knowledge (Berger and Luckmann 1966) neglected media, and so is prima facie hardly a useful reference-point for attempts to grasp our embeddedness in a digital world. However, phenomenology's project of excavating the many layers of taken-for-grantedness in everyday action remains a provocative starting-point for grasping the lived tensions of a world of 'computed sociality' (Kallinikos and Tempini 2014): it is worth asking, for example, to what extent Schutz's statement that 'not only the what but also the how of the individual situation in the lifeworld belongs to the fundamental elements of the stock of knowledge' (1973: 105)

applies to the era of algorithmic power, and what follows, to the extent that it does not. We can only get at this however if we start listening to people's attempts to manage 'the how' of their 'situation in the lifeworld' under digital conditions. This is what the project of social analytics seeks to do.

As Star and Ruhleder (1996) have noted, a key dimension of an infrastructure is its embeddedness, in other words how it 'sinks' inside social arrangements, and how it shapes and is shaped by the conventions of a community of practice (Lave and Wenger 1992). To study the development of a digital infrastructure that 'sinks' into the background of social aims and community practices is not, as we discuss later, without challenges. The gap between infrastructure and enactment is *all the more* important, when actors choose to take an active role in thinking about the analytics they use, how they are designed, and the online information architectures which they help generate.

Three methodological points flow from this distinctive attention to reflection and agency. *Firstly*, we emphasise *agency* in relation to social analytics as a performative process that develops *over time*. Algorithmic power does not act in the abstract; it must be materialised in various ways, and become embedded in processes of action, adjustment, recalibration, and stabilization. Such processes take time; they may change *over* time. Our fieldwork involved the long-term monitoring of practice (over 18 months): it registered emerging 'resistance' within the intertwining of human and material agency (Pickering 1995: 65-67). One type of resistance noted by Bowker (2005: 116) stems from an everyday instrumentalism, whereby people resist spending additional time on creating and preserving information *about* their data. *Secondly*, following the core principles of action research, our social analytics approach worked *collaboratively*, implementing analytic techniques in evolving situations with a

partner organization. The partner discussed in this article was affected, like any other, by overlapping forms of power relations and resource constraint, which affected our research outcomes. *Thirdly*, we approached this research within a recurring cycle of action *and reflection*. We were interested in facilitating a process where the conditions for agency in an analytics-saturated world could be increasingly understood by social actors, but also how they then might be acted upon. Our interest in agency and in the tensions, frictions and blockages that limit agency, required translating technical issues into non-technical terms in order to encourage reflection by our partner and, in turn, by us on the actions taken subsequently by our partner. Social analytics research extends digital social research (Marres and Weltevrede 2013) into concrete contexts of interaction, making their phenomenological complexity all the more explicit.

Doing Social Analytics Research

Real social actors may encounter analytics in situations where the encounter is not already regulated by clear goals and rules. The encounter may be messier, particularly where those actors' institutional aims are not already aligned with outcomes directly related to those analytics, as for the community organization with which we conducted our fieldwork.

Our fieldwork partner was an organization whose goal was to strengthen community voice; to do so, they knew they needed to enhance their digital presence, including via better use of analytics. When we first talked to them, they had a website which displayed various outputs of community voice, but received little traffic. It was through thinking about how we could help them implement their aspiration that we formulated 'social analytics' research as a more general programme. Working with this partner brought us face-to-face with what Pickering

(1995: xi) calls ‘resistance’, that is, emergent mismatches between human and non-human agency.

Overall research design

Our partner (which to preserve anonymity we call ‘C-Media’) was situated within the civil society sector. It was clear about the importance of online networking as a key means of producing community voice, and had aims for national profile, even though its main activities were in one particular large city in the North of England. In its concern to promote individually and collectively produced content C-Media had similarities with many contemporary arts, social and community organizations, and overlaps with other organizations for whom content production and circulation is a subsidiary aim. For this reason, we believe that our fieldwork with C-Media has value as an exemplar of understanding the actions of many organizational actors.

Our wider research aimed to study not C-Media’s use of analytics as such, but rather how its use of digital infrastructure supported its wider *social goals*. We started with a review of C-Media’s existing practice and its interrelations, especially via its website, with its community of content producers and the content they produced. We wanted to know how C-Media understood this interrelation, and how it was supported by its website’s existing operations. When we first began working with C-Media, their site aggregated content, but with a minimal attempt to ‘curate’; there was also minimal use of data collection and external algorithms such as those used by Google and Facebook.

Our research collaboration responded to two factors. *First*, C-Media wanted to use web analytics to gain a sense of how their website was being used and to begin collecting basic headline data from Google Analytics. It became clear that they had 'answers' (that is, outcomes they wanted to achieve and support through their digital presence) but as yet no well-formulated 'questions' (that is, no structured means of thinking about their digital presence, and what online practices might enable them to enhance it in these specific ways). In response, we started with C-Media considering explicitly: what 'questions' could be asked through appropriately adapted and customised 'analytics', how could the data received be interpreted so as to generate rigorous 'answers', and what would be the limits of such answers in the context of their practice? Our aim was to initiate a collaborative process of 'thinking through' analytics, not in the abstract, but in relation to C-Media's social aims. This process fitted into what C-Media itself indicated it wanted from the research. *Secondly*, C-Media had built their content site to include many of the systems discussed in literature on algorithms and relational databases, for example user profiles, metadata for searching content, and rating systems for content. They already had some awareness of how to restructure their online practice in a way that was responsive to analytic data. But while they used these basic tools of web 2.0, they lacked the means for formulating specific *actions in response* to generated data, and especially a way of thinking practically about their potential for shaping interactions at the 'front end' of their site to better meet their overall social goals. The basic tools of analytics and the aims they served were therefore present, as was the desire to further those aims in the wider service of C-Media's social goals, but they were not embedded in C-Media's everyday practice.

Our challenge therefore was to help C-Media identify digital outcomes that would satisfy its broader goals (eg more site traffic beyond the audience of its content producers ('CPs'); more

interactive engagement with the stories uploaded; more evidence that these stories, or pieces of content, helped to sustain a CP network) *and* practical means for achieving those outcomes from changes in its website's basic architecture and interface to its metadata. Reframing this in phenomenological terms, our aim was to help C-Media 'know itself' by making its digital presence an effective object of reflection and action. We agreed with C-Media that this would involve, first, developing an explicit *framework* for systematically collecting, organising, interpreting and visualising data directed towards achieving clear organisational outcomes; second, *tracking* C-Media's internal reflections on the collection and interpretation of that data and its monitoring and analysis, as part of C-Media's working culture, with our field comments working as a 'mirror' in which C-Media's deliberations were reflected back to itself; and third, supporting C-Media in the design and *implementation* of this system through meetings with key staff over a sustained period, while simultaneously *monitoring* this process.

This three-stage approach (framing; tracking; monitored implementation) is, we propose, necessary to doing social analytics research in practice. Central to it was the use of certain digital tools.

Initial Steps

As part of our broader collaboration with C-Media, we discussed what measurements would serve as indicators for how their website was serving their social goals. We helped re-design its training offer and accreditation structure in ways that generated data that could be measured in the form of such indicators. Developing the digital infrastructure for that data collection was a precondition for the realisation of C-Media's social aims.

This preparation process also entailed helping C-Media adjust their administrative processes so that they could, in the future, support the web analytics provided by the platforms. Drawing on earlier experience of one member of our research team in building sustainable skills within civil society groups, we took an 'outcomes' approach.[5] Two intended outcomes of this framework are particularly relevant here: to adjust the types of content produced in ways that could help C-Media better meet its social aims; and to manage begter this content so as to enhance public engagement with it[6].

Measurements

Our initial focus was on measuring the quantity of uploaded material to the website and the instances of specific tags used by CPs to characterise their content. These measurements helped us get a sense of how (or if) the dominant themes on the website reflected the social goals of the organisation. In addition, and in order to enhance the levels of active audience engagement with the content online, it was important to map the different constituencies and publics forming around the website. Here, we were less interested in audience numbers or website hits, and more on how visitors initially found their way to the site. For instance, where were web visitors located as they visited the site, what was their interest and how did they interact with content and with each other?

For the collection of these data we considered, with C-Media, both web analytics services which maintain some access to the data collected such as Google Analytics (GA), and open source analytics such as Drupal[7] which allow more flexible processes for analysing and combining data about one's own site. Eventually, a combined usage of open source and proprietary platforms was developed which allowed the collection of statistics about the different search terms that visitors used upon landing on the website. This helped to

differentiate between visitors looking for the site and those reaching the website when searching for something unconnected.

In this way, C-Media would be able to observe reflexively whether their efforts in opening up to new audiences had been successful and, as discussed next, to revise the ways in which their content was categorised and tagged.

Location and engagement with content

In addition to information about how they reached content on the website, the analytics collected allowed C-Media to locate visitors geographically. For example in July 2012, most visitors accessed the website from C-Media's local area, and from areas where C-Media's programme was delivered. This information was important given C-Media's national and potentially international networking ambitions.

In addition to data about visitor traffic and visitor location, it was important for C-Media to collect and analyse information about how these visitors engaged with the site, in order to meet its social aim of increased public engagement. We measured the number and content of comments on the website, and the ways in which content circulated in social media. These data allowed C-Media to observe social media activity, such as social connections and exchanges via Facebook or email, and behaviours such as 'liking' (at that stage rather low). With access to that information, C-Media could rethink not only how they facilitated standard 'social media' engagement in the website design, but also whether that was a broader type of engagement which they wanted to facilitate. C-Media resisted being 'locked into' the prescribed functions offered by social media platforms and instead tried to control actively how reporters were able to interact with one another via the site. Over time, C-Media chose to

encourage engagement using the comments function (an existing feature on C-Media's site), but they also decided to introduce new ways of user engagement, such as a module which allowed users to rate and collectively categorise content, through the Drupal modules 'Five Star' and 'Community Tags'.

Content curation experiments

Basic measurement and information-gathering was not enough to make this a social analytics project. It was important also to track C-Media as they started to *reflect* on how their website was serving their wider social aims. We were able to agree with C-Media a set of outcomes (presented in a wider document whose underlying form is adaptable to a range of contexts [**link to document on project website**]. Important also was planning with C-Media ways in which standard analytic measures (e.g. visitor distribution, engagement with content) could be *put into action* in defined contexts that would test if C-Media's organizational purposes were being supported or not, providing a basis for further reflection and adjustments. We followed these steps of inquiry, planning, reflection and readjustment (the four stages of an action research cycle: Foulger 2010), and after the initial collection of metadata, we proceeded with the digital adjustments of the website. These adjustments took the form of content curation experiments[8].

The content curation experiments drew various people linked with C-Media into thinking about new ways of working with digital content and engaging audiences in digital interaction. Initially, we had found that the limited functionality of the website did not enable users to connect to others, or curate content themselves. Throughout the winter of 2012-2013, while the website was being re-designed, we reflected with C-Media's Board on the implications and practicalities of content curation. An important part of the expectation was for users, and

not only full-time staff, to learn how to 'tag' their articles before publishing them online and how to use other kinds of metadata in order to have their articles appear in search engines and so reach readers beyond C-Media's immediate network.

The basic web measures (visitor distribution, engagement with content) which C-Media had already started to collect with the use of GA and Drupal, were essential for this experiment because they manifested how C-Media's news stories, even in CP blogs, had a remarkably short 'life cycle'. Having evidence (from collected metadata) that, for instance, a story about violence in their community lost its appeal after a week, made C-Media consider ways to increase the longevity of their stories. Thinking about measurements that could indicate the life of a story brought C-Media face-to-face with other needs: to update their content regularly; to invest in reporting specific communities' news, as opposed to acting as a general platform.

These reflections fed into a more detailed plan for curation experiments which built structured reflection into C-Media's organizational rhythms. C-Media initiated a tentative twelve-month plan of guest curators who would each focus on a specific theme, relevant to their experience and expertise. Good examples of this practice came from online artists' agencies, which feature curated content from their collections. Eventually, these curators were decided by an Editorial Board or Steering Committee, which brought together CPs from Germany, Italy and Sweden. In addition, and in recognition of the need for sustained editorial support, C-Media contemplated introducing a new editorial role responsible for curation experiments. Yet a lack of resource blocked the creation of this new editorial role.

Socially generated metadata

If ‘social analytics’ is interested in social actors’ reflexivity about their interfaces with digital infrastructure, then it must encompass the possibility that analytics’ use can be involve a wider range of actors than just the full-time staff at the organization whose data is being measured. For C-Media this meant recruiting and mobilising the wider community of content producers focussed around C-Media’s practice. This led to the idea developed with C-Media of setting up on its site the functionality for *user-generated* meta-data: *community tagging*. Social or user tags, often also referred to as ‘folksonomies’ and social bookmarking, are a type of metadata in the form of keyword descriptions. They are often used in blogs, and increasingly in other web 2.0 application, to identify content (image, text) within a site. They allow the indexing and wider categorising of webpages with identical tags, which consequently enables users to search for similar or related content. ‘Community tags’ is a module available in Drupal (the open-source platform used by C-Media), which allowed CPs to tag content, and track who tagged what and when: in other words, a *user-generated* tagging system. The aim was both to curate new content and to deepen user engagement with existing content. By actively engaging audiences in the selection of material which would be featured on the website, we hoped to encourage a culture of sustained discussion and an awareness of the significance of metadata management. As the chief executive of the organisation noted before the experiments started:

> What I’d like on the front page is something which talks about the curated content stories that we come up with, […] featuring stories and we can have guest bloggers for that, but also by a natural process of almost number of hits appears on the front page. And I love the idea of the community tagging […] because that supports an idea that we’ve got around mass collaboration that people themselves will self-select the content and create the stories that they’re evolving. (interview 16/10/2012).

As time passed and the first experiments were completed, C-Media's management became increasingly enthusiastic about content curation's potential to contribute to C-Media's digital profile:

> So for me that thing about raising the profile means really thinking about the audience slightly differently and wanting to target those agencies, organisations or people that might take the idea of the movement [of community reporters] forward and help it develop on our behalf, so for me that was a real shift in my thinking about who the website was for and what the purpose of it was. I still want it to be somewhere that community reporters can go and hang out and develop relationships and skills virtually, I'm still really keen for that, but what the content does and who it talks to, it's got a job to do, I don't want it just to be an exercise in vanity for the reporters that they're online, I want it to work a lot harder than that. (interview 16/10/2012).

The community tagging experiment we performed with C-Media embedded algorithms within the Drupal 'community Tags' module (such as 'Tag of the Week') into C-Media's home page, inviting CPs to categorise and tag the content (so helping the organisation make this content more 'searchable' by others). From this starting-point, tags made could be aggregated and counted, partly through automated algorithmic processes. The result was a tag cloud of all tags in a vocabulary, displayed on C-Media's front page. In addition to the community tagging system, C-Media provided guidance for first time users of its website looking to upload material for subsequent tagging and searching, and planned to rate content via the home page, in a section called 'Most Viewed Stories'.

Longer-term, we discussed with C-Media using platforms such as Storify (https://storify.com) in order to draw together multi-media web content around a featured theme (additional commentary, content), and to experiment with multiple story authorship. But such enhanced public engagement with their content (and the resulting complication of their digital presence) required of C-Media a systematic and long term investment in the development of skills and new 'data literacies' (van Dijck 2013): in other words, more resource than was available at this stage of the project.

We have outlined in this subsection a cumulative series of actions, reflections and digital tool development undertaken collaboratively with C-Media which, taken together, constitute a basic model for conducting social analytics research with any social or community partner. A precondition is that the partner has social goals which it recognises can be enhanced by the refinement of its digital presence. The translation of the goals into practice, and the wider reflections such translation then generates, are the core on which social analytics seeks to focus.

Discussion and Conclusion

This article began with a phenomenological challenge: if 'presence' for social actors now necessarily involves being categorised and measured by automated processes over which the social actor may have little control, or even awareness, how is this relationship experienced? Going beyond general diagnoses of algorithmic power, this article has offered a distinctive project of empirical research: social analytics.

Social analytics tracks in everyday settings how basic analytical tools are used by social actors to fulfil their *social* ends. This involves understanding analytics as encompassing not

merely the basic tools of measurement now routine in most contexts of online interaction, but also the wider operations and reflexive adjustments on which social actors' digital presence now relies. A project of *social* analytics foregrounds the lived tensions and contested reflexivity involved when a social actor seeks to maintain its presence in the world under these new conditions. There are few tensions, of course, for social actors whose *goal* is to develop and apply analytics, or directly to profit from the data they generate, even if there may be many practical complexities (Gillespie 2014 on Twitter). 'Social anlytics' as a sociological project is not, however, concerned with such actors, but with actors whose goals are themselves specifically social, such as the community-based organization where the project evolved during fieldwork. Such social actors, far from being operating unreflexively under the weight of algorithmic power, are now required to *pursue* reflexivity about their goals under 'algorithmic' conditions not of their own choosing, but over which they may seek to acquire some reflexive control. Tracking those processes of reflexive adjustment, in all their complexity, distinguishes social analytics' concerns from the everyday use of analytics *as such*, which is now routine in countless organizations, enterprises and businesses. Indeed, from the study of such reflexive tensions emerges a new topic for sociology and for a phenomenology of the digital age, which goes unnoticed by general analyses of algorithmic power.

The promise of 'social analytics' for studying a range of social actors has emerged through the details of the case study presented here. Seen from the perspective of social analytics, apparently banal questions of website design, metadata selection, interaction tracking and the means for categorizing and linking information about online interactions become a mode in which, sociologically, we can track a social actor's reflections about itself, and its possibilities for sustaining a particular presence – that is, a presence aligned to its distinctive

social ends – through better management of the operations of the digital interfaces by which it has no choice but to be judged. Within the case study's detail, we saw emerge reflections for example on how C-Media's website should be monitored, on the basis of what types of categorization, and who (beyond C-Media's managers) should be involved in developing such categories ('community tagging'). The means were, in a sense, technical, but the goals were not (for example, to increase the longevity of story content): the means too (for example, exercises in content curation) were often much more than technical devices. Indeed, under these conditions, the distinction between 'social' and 'technical' is not useful. For C-Media, making its content 'work' (for its broader goals) was inseparable from reflecting on how the mechanisms of its digital presence operated in alignment with those goals. Phenomenologically, a social actor's presence is now automatically implicated in such operational details, generating new issues of *translation* between automated processes of measurement and the broader formulations inherent in social goals.

As a result, social analytics must also track tensions and resistances. Changes in an organization's reflective process, especially at a time of acute resource constraint, require a long time-period: even 18 months was not long enough to see through the process we had proposed for implementation. Funding limitations restricted staff resourcing and time, and so limited actual fieldwork processes: for example, synchronising existing reporter web accounts with all traces of their presence online (as required for the planned mapping process) proved too much for the single web designer that could be employed. A key administrator left C-Media during the fieldwork; the collection of analytics was interrupted, and only resumed after we had left the field. But such difficulties could affect *any* fieldwork that seeks to investigate social analytics in a collaborative setting.

There were also at times disagreements within C-Media about the detailed aims of their digital development. During web development meetings, there were long discussions as to whether content produced during proposed content curation experiments should aim to cover timely events, or focus on these events' aftermath. Such discussions revealed deep-rooted differences in how individuals understood the content sector in which they operated. C-Media as an organisation, and its CPs as individual creators, were being asked to step back from everyday practices of content creation and focus on the deeper purposes for which they were trying to manage content *at all.* Conflicts between the basic 'information' produced by basic analytics and the wider meanings of those measures for that organization's social aims are inherent, not incidental, to researching social analytics and to the phenomenology of a digital world.

Underlying these tensions is the point that, under conditions where 'social' (or community) aims cannot be separated from 'technical' implementation and the 'technical' is already deeply cultural (and social), basic analytics were both *close to* and *remote from* C-Media's everyday practice. For sure, analytics may seem close to the aims of a social actor (because, for example, its funding body's targets make specific reference to analytic-based measures), but their *implementation* still involves considerable new learning, which requires resource that may not be to hand. Social analytics research must take account of such everyday resource-based 'resistance' (Pickering 1995: xi) to processes of measurement.

The research project of social analytics emerged unexpectedly from broader fieldwork. In future fieldwork, we anticipate working with social actors *from the start* around their *explicit* aim of reevaluating their digital presence, and its fit with their social civic or cultural ends. The elaboration of such explicit aims will depend on the organization studied: there are many

possibilities, across the arts and cultural sectors, civil society, and social and community organizations. Importantly, such actors need not have aims that relate explicitly to 'information politics' (Rogers 2004), since it is the complexity of translating a wide range of social goals *into* the operations of digital infrastructure with which social analytics is concerned.

When the 'self' of social actors (institutional or individual) is necessarily translated through automated processes of measurement, operating via the 'back-end' of the digital world (Rogers 2004: 3), social actors' attempts to maintain some reflexive control of that process are a core site of social contestation and a core topic for social phenomenology. The aim of social analytics is to recover this important aspect of contemporary everyday practice for sociological reflection.

References

XXXXa

XXXXa

XXXXb

XXXXb and XXXXa

Baym, N. 2013 'Data Not Seen: The uses and shortcomings of social media metrics', *first Monday* 18(10), 7 October. http://firstmonday.org

Beer, D. 2009 'Power through the Algorithm? Participatory web cultures and the technological unconscious', *New Media & Society* 11: 985-1002.

Beniger, J. 1986 *The Control Revolution.* Cambridge, Mass.: Harvard University Press.

Berger, P. and Luckmann, T. 1966 *The Social Construction of Reality*. Harmondsworth: Penguin.

Berry, D. 2011 *The Philosophy of Software.* Basingstoke: Palgrave MacMillan.

boyd, d. and Crawford, K. 2011 'Critical Questions for Big Data: Provocations for a cultural, technological and scholarly phenomenon', *Information, Communication and Society.* 15(5): 662-679.

Boellstorff, Tom. 2008. Coming of Age in Second Life: An Anthropologist Explores the Virtually Human. Princeton, NJ: Princeton University Press.

boyd, d. (2014). It's Complicated: The Social Lives of Networked Teens. New Haven: Yale University Press.

Bowker, G. (2005) *Memory Practices in the Sciences.* Cambridge, Mass.: MIT Press.

Bowker, G. and Star, S. Leigh (1999) *Sorting Things Out.* Cambridge, Mass.: MIT Press.

Brighenti, A. 2007 'Visibility: a category for the social sciences', *Current Sociology* 55(3): 323-342.

Burrows, R. 2009 'Afterword: Urban Informatics and Social Ontology' in M. Foth (ed.) *Handbook of Research on Urban Informatics*, Hershey, PA: Information Science Reference, 450-454.

Chandler, A. 1977 *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, Mass.: Belknap Press.

Clanchy, M. 1992 *From Memory to Written Record: England 1066-1307.* 2nd edition. Oxford: Blackwell.

Cooley, C. 1902 *Human Nature and the Social Order*. New York: Charles Scribner's Sons.

Crossley, N. 2001 *The Social Body* . London: Sage

Elias, N. 1994 *The Civilizing Process*. Oxford: Blackwell.

Ellison, N., Heino, R., and Gibbs, J. (2006). 'Managing Impressions Online: Self-presentation. Processes in the Online Dating Environment'. *Journal of Computer-Mediated Communication*,11/2

Foulger, T. 2010 External Conversations: An unexpected discovery about the critical friend in action research inquiries. *Action Research.* 8, 135-152.

Fuller, M. 2009 *A Software Lexicon.* Cambridge: MIT Press.

Gerlitz, C. and Helmond, A. 2013 'The Like Economy: Social buttons and the data-intensive web', *New Media & Society.* first published on February 4, 2013 doi:10.1177/1461444812472322

Gillespie, T. 2014 'The Relevance of Algorithms' in P. Boczkowski, Foot, K. and Gillespie, T. *Media Technologies.* Cambridge, Mass.: MIT Press.

New Haven: Yale University Press.

Greenhalgh, T., Stones, R. and Swinglehurst, D. 2014 ‚Choose and Book: A sociological analysis of „resistance" to an expert system', *Social Science & Medicine* 104: 210-219.

Hacking, I. 1999 *The Social Construction of What?* Cambridge, Mass.: Harvard University Press.

Halavais, A. 2009 *Search Engine Society*. Cambridge: Polity.

Halford, S., Pope, C. and Weal, M. 2013 'Digital Futures? Sociological Challenges and Opportunities in the Emergent Semantic Web', *Sociology* 47(1): 173-189.

Halford, S. and Savage, M. 2010 'Reconceptualizing Digital Social Inequality', *Information communication & Society* 13(7): 937-955.

Heinich, N. (2012) *De La Visibilité*. Paris: Gallimard.

Hindman, M. 2009 *The Myth of Digital Democracy.* Princeton: Princeton University Press.

Introna, L. 2011 'The Enframing of Code: Agency, Originality and the Plagiarist', *Theory Cutlure & Society* 28(6): 113-141.

Introna, L. and Nissenbaum, H. 2000 'Shaping the Web: Why the Politics of Search Engines Matters', *The Information Society* 16: 169-185.

Kallinikos, J. 2009 'The Making of Ephemeria: On the Shortening Life spans of Information', *The International Journal of Interdisciplinary social sciences,* 4(3): 227-236.

Kallinikos, J. and Tempini, N. (2014) 'Social Data as Medical Facts: Web-based Practices of Expert Knowledge Creation', *Information systems Research,* 25(4): 817-833.

Kling, R. 2007 'What is social Informatics and why does it Matter?' *The Information Society* 23(4): 205-220.

Lash, S. 2002 *Critique of Information*. London: Sage.

Lash, S. 2006 'Dialectic of Information'? A Response to Taylor', *Information Community & Society* 9(5): 572-581.

Lash, S. 2007 'Power after Hegemony: Cultural Studies in Mutation', *Theory Culture & Society* 24(3): 55-78.

MacKenzie, A. 2006 *Cutting Code.* New York: Peter Lang.

MacKenzie, D. and Wajcman, J. (eds.)(1999) *The Social Shaping of Technology.* Milton Keynes: Open University Press.

Marres, Noortje and Weltevrede, Esther. 2013. Scraping the Social? Issues in live social research. *Journal of Cultural Economy*, 6(3), pp. 313-335. ISSN 1753-0350

Orlikowski, W.J. and Scott S.V. 2014 'The Algorithm and the Crowd: Considering the Materiality of Service Innovation' *MIS Quarterly*, Special Issue on Service Innovation. http://www.misq.org/skin/frontend/default/misq/pdf/Abstracts/OrlikowskiScott_SI11663.pdf

Pickering, A. 1995 *The Mangle of Practice*. Chicago: Chicago University Press.

Postill, J. 2013 'Democracy in an age of viral reality: A media epidemmiography of Spain's indignados movement', *Ethnography,* first published on October 23 2013 doi: 10.1177/1466138113502513.

Power, M. 1997 *The Audit Society.* Oxford: Oxford University Press.

Rogers, R. 2004 *Information Politics on the Web*. Cambridge, Mass.: MIT Press.

Rogers, R. 2010 'Internet Research: The Question of Method', *Journal of Information Technology and Politics* 7: 241-260.

Rogers, R. 2013 *Digital Methods*. Cambridge, Mass.: MIT Press.

Scott, S. and Orlikowski, W. 2013 'Sociomateriality – taking the wrong turning? A response to Mutch', *Information and Organization* 23: 77-80.

Simmel, G. 1990 *The Philosophy of Money*. London: Routledge.

Slater, D. and Miller, D. 2000 *The Internet: An Ethnographic Approach.* Oxford: Berg.

Suchman, L. 2007 *Human-Machine Configurations.* Cambridge: Cambridge University Press.

Turkle, S. 2011 *Alone Together.* New York: Basic Books.

Turow, J. 2007 *Niche Envy*. Cambridge, Mass: MIT Press.

Turow, J. 2011 *The Daily You.* New Haven: Yale University Press.

Van Dijck, J. 2013 *The Culture of Connectivity.* Oxford: Oxford University Press.

Van Dijck, J. 2014 forthcoming 'Datafication, Dataism and Dataveillance: Big Data between scientific paradigm and ideology', *surveillance and Society* 14(2):

Yates, J. 1989 *Control through Communication: the Rise of System in American Management.* Baltimore, MD: Johns Hopkins University Press.

---

[1] http://www.webanalyticsassociation.org/?page=aboutus

[2] Note on funding

[3] In medical sociology, see Greenhalgh, Stones and Swinglehurst (2014) on resistance to the introduction of technological expert systems into medical practice; in management studies work, see Orlikowski and Scott (2014) on the tourism industry; in political sociology, see Postill (2013) on Spanish protest groups reflexive use of the trending features of Twitter.

[4] The impact of the May 2014 EU court decision against Google Spain on this overall situation remains unclear.

[5] An ‘outcomes’ approach involves assessing the changes that programmed activity is intended to deliver on, and is a common means of project self-evaluation within public, third sector and social enterprise organisations: see our sample Outcomes Framework documents [url].

[6] Other planned outcomes, which did not involve collecting metadata or web analytics, were logging formal training and individual skills development, sustaining local group interactions through informal ‘meet-ups’, and networking.

[7] Here we refer to the ‘Views’ template in the Drupal platform.

[8] A detailed account of this aspect of our research is given in forthcoming articles (XXXXa and XXXXb)