



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/152596/>

Version: Accepted Version

Article:

Gerrard, Y. and Bates, J. (2019) Introduction to the data power special issue : tactics, access and shaping. *Online Information Review*, 43 (6). pp. 945-951. ISSN: 1468-4527

<https://doi.org/10.1108/oir-10-2019-396>

© 2019 Emerald Publishing. This is an author-produced version of a paper subsequently published in *Online Information Review*. This version is distributed under the terms of the Creative Commons Attribution-NonCommercial Licence (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. You may not use the material for commercial purposes.

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

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



INTRODUCTION TO THE DATA POWER SPECIAL ISSUE: TACTICS, ACCESS, & SHAPING.

The articles in this *Online Information Review* (OIR) Special Issue were presented at the Data Power Conference 2017¹ (Carleton University, Canada, 22-23 June), organised by local hosts Dr Tracey Lauriault and Dr Merlyna Lim, with support from the wider Data Power steering committee Prof. Helen Kennedy, Dr Jo Bates and Dr Ysabel Gerrard (Sheffield, UK).

Now approaching its third iteration², the Data Power Conference focuses on critical questions about the relationship between data and society, with conference speakers invited to address the social and cultural consequences of data's pervasiveness in our everyday lives. With this focus on emergent 'data relations' (Kennedy, 2016), the Data Power Special Issue brings a different slant to the advance of datafication and algorithmic processing than is commonly seen in regular issues of OIR. Papers were selected by the full conference team for their quality, as well as their relevance to the Information Science research community who make up the majority of the OIR readership.

The Data Power Conference 2017, and by extension the seven papers in this special issue, addressed three questions:

¹ The conference was supported by the following funders: The Social Sciences and Humanities Research Council of Canada Connection Programme 2017; Faculty of Public Affairs, Carleton University; School of Journalism and Communication Carleton University; MacOdrum Library, Carleton University; Carleton Institute for Data Science (CUIDS); Carleton Institute of Criminology and Criminal Justice; Individual Research Grant contributions from Tracey P. Lauriault and Jeffrey Monahan, Carleton University; and from Merlyna Lim's Canada Research Chair Grant.

² The next Data Power Conference will be held at the University of Bremen, 12-13 September 2019.

1. How can we reclaim some form of data-based power and autonomy, and advance data-based technological citizenship, while living in regimes of data power?
2. Is it possible to regain agency and mobilise data for the common good? To do so, which theories help to interrogate and make sense of the operations of data power?
3. What kind of design frameworks are needed to build and deploy data-based technologies with values and ethics that are equitable and fair? How can big data be mobilised to improve how we live, beyond notions of efficiency and innovation?

These questions broadly emphasise the reclamation of power, retention of agency, and ethics of data-based technologies, and they reflect a broader moment in recent data studies scholarship. While early critical research on ‘big data’ – a term that captures the technologies, analytics, and mythologies of increasingly large data sets (boyd and Crawford, 2012) – could only hypothesise the inequalities and deepened forms discrimination that might emerge as data sets grew in volume, many of those predictions have now become real. The articles in this Special Issue ask pressing questions about data power at a time when we have learned that data are too frequently handled in a way that deepens social inequalities and injustices (amongst others, Eubanks, 2018; Noble, 2018).

The papers in this Special Issue approach discussions of inequality and injustice through three broad lenses: (1) the *tactics* people use to confront unequal distributions of (data) power; (2) the *access* to data that is most relevant and essential for particular social groups, coupled with the changing and uncertain legalities of data access; and, (3) the *shaping* of social relations by and through data, whether through the demands placed on app users to disclose more personal information, the use of data to construct cultures of compliance, or through the very methodologies commonly used to organise and label information. While these three themes do not exhaustively capture the range of topics addressed in this Special Issue, at the Data Power Conferences, or within the field at large, they represent an

emphasis within data studies scholarship on shedding light on the most pressing issues confronting our increasingly datafied world.

PART ONE: TACTICS.

Two of the papers in this Special Issue – ‘Datafication, dataveillance, and the social credit system as China’s new normal’ (Lee, 2019) and ‘What difference does data make? Data management and social change (Paris et al, 2019) – offer unique critiques of power relations through the framework of ‘tactics’. It is worth engaging in a brief discussion of Michel de Certeau’s *The Practice of Everyday Life* (1988) here to frame these contributions. This book explores what de Certeau calls ‘making do’: the tactics people use to navigate power structures in their everyday lives. De Certeau (1988) talks about the everyday, commonplace power struggles between two social groups: the producers of culture and their users (ordinary people), and argues that people ‘conform’ to mechanisms of discipline and power ‘only in order to evade them’ (1984, p.xiv). In other words, it is possible for people to evade/subvert the ruling order by using a set of *tactics*. But as de Certeau notes, ‘Whatever it wins, it does not keep. It must constantly manipulate events in order to turn them into “opportunities”’ (de Certeau, 1984, p.xix). The necessity to constantly evolve one’s tactics is a particularly important consideration for Lee (2019) and Paris et al (2019), who write about the tactics people are adapting during the current climate of fast-paced data-based changes (indeed, at a pace that de Certeau probably did not foresee).

Opening this Special Issue, Claire Lee’s paper shares novel early insights into how Chinese citizens adopt tactics in the face of the social credit system: a form of societal governance that intends to standardise Chinese citizens’ behaviour and reputation by collecting personal information to evaluate citizens and give them a ‘social credit score’. As Lee explains:

With a high score, one can easily acquire a cash advance, obtain expedited access to vital services including medical insurance and education, as well as enjoy faster processing at customs, in applying for visas and securing animal adoptions. On the other hand, individuals who have lower scores will be restricted from services. (2019: X)

The social credit system is still in the early phases of its development, but the plan has been a long-term goal for the Chinese government and citizens are becoming more aware of its steady implementation. Lee collected personal narratives from Chinese citizens to learn their views on what some refer to as ‘the new normal’ in Chinese society. Lee’s main research finding is that the social credit system has heightened citizens’ practices of self-surveillance (see also Lupton and Williamson, 2017): Lee’s participants *feel* as though they have begun to monitor themselves more closely since they learned about the social credit system and its consequences. Lee concludes the paper by raising a point of discussion, asking how Chinese citizens might ‘disrupt the system from the inside’ (2019: X), and wondering which citizens will be disproportionately affected by the social credit system. Lee puts herself in conversation with authors whose work has recurred across this Special Issue and who focus primarily on the consequences of data power, such as boyd and Crawford (2012), Lyon (2014), and O’Neil (2016), amongst many others.

Britt Paris, Morgan Currie, and Joan Donovan then turn to a discussion of the data management practices undertaken by activist groups and grassroots organisations. The authors expand on emerging data activism literature to draw distinctions between the data infrastructures used by groups that organise in response to data collection by corporations and the state. For example, they explore how Fatal Encounters used a collective database to produce missing data about police homicides in the US, offering a critique of the transparency of publicly available data sets. They examine how Making Sense created a new dataset to disclose information about the air quality around Kosovo, an issue that had been concealed by their government; and, how DataRescue – led by the Environmental Data

Governance Initiative (EDGI) – archived data created by US federal scientists that documented evidence of climate change and human-induced ecological violence. Literature on data activism typically focuses on generating new data, but the case studies shared by Paris et al (2019) demonstrate how activist groups use missing or already-existing data to highlight contentious political issues. Data management systems are thus a crucial part of political mobilisation for these organizations: without proper management, activist groups would simply be unable to do this work. The authors argue that the results activists produce are only a small part of the work they do: ‘Behind the visualisations or public-facing databases are a suite of data management infrastructures and organisational norms that form a considerable part of activists’ mundane practice’ (Paris et al, 2019: X). As Paris et al note, data infrastructures are academically under-addressed and yet are integral to ‘shaping the tactics and political formation of data activists’ (2019: X). While Lee and Paris et al’s focus is on the tactics citizens use to respond to different contexts of datafication, our next two papers draw attention to issues around citizens’ access to data relevant to their needs.

PART TWO: ACCESS.

Two of the papers in this Special Issue – ‘Ownership and control over publicly accessible platform data’ (Scassa, 2019) and ‘Open government for all? Co-creating digital public services for older adults through data walks’ (Jarke, 2019) – offer critiques on the availability and scope of online data sets. Although all of the papers in this Special Issue deal in some way with data and inequality, the papers in this section address issues of data access for social groups and actors whose data needs are often marginalised within discourses and practices dominated by powerful institutions. Readers are likely familiar with such struggles in the context of the increasingly restricted access to social media platforms’ Application Programming Interfaces (APIs) used by researchers to collect data and produce knowledge about the social world. Our authors examine similar issues in the context of the legalities of

access to Airbnb data, and in the development of methodological approaches to make open government data driven services more relevant to the needs and experiences of older citizens.

Using Airbnb as a case study, Teresa Scassa explores the legalities of access to publicly available data in what van Dijck et al (2018) and others have called the 'platform society'. In the paper, Scassa explains that a diverse range of parties make use of publicly available Airbnb data for reasons which often serve the public interest, but existing legal frameworks are not particularly well-suited to 'our evolving data society generally or to platform data ecosystems in particular' (2019: X). This creates a risk that users' perspectives and the public interest will not be well-represented in evolving litigation, if indeed 'they are represented at all' (Scassa, 2019: X). Scassa argues that Airbnb's data has significance beyond its immediate user base because we can learn about, for example, 'the platform's effects on the cost and availability of long term accommodation, its impact on incumbent short-term accommodation providers, the incidence of discrimination in Airbnb rentals and pricing' (2019: X). Although there is a strong public need for access to data by companies like Airbnb, there remain legal uncertainties in relation to the ownership and rights of access to publicly accessible data. Scassa reminds readers that the legitimacy of data scraping activities are likely to be decided by litigation between large, wealthy commercial competitors because, simply put, litigation is expensive. Perhaps the biggest risk is that access to data will be resolved by litigation between business competitors, and will therefore not bear wider interests in mind. This risks squeezing out the voices and interests of non-commercial users, along with non-profit organizations and smaller companies. While we might not all agree on which kinds of data serve the public interest, Scassa reminds us that stark inequalities will emerge if data's accessibility is determined only by private interests.

Juliane Jarke's article offers an example of data related discrimination through the lens of Open Government Data initiatives. As Jarke explains, Open Government Data initiatives promote the ideals

of 'citizen collaboration and participation in the planning, design and delivery of public services' (2019: X):

The idea of using open government data for new public services is simple: Governments provide their data for free, online and under open licences; civil society actors or private companies may re-use the data and develop services according to their needs or expected demand. (2019: X)

One of the main espoused benefits of Open Government Data is that it can foster the development of user-centered services by third-parties without incurring additional costs on the state. Ideally, this means citizens should receive better services, however Jarke observes that the needs of older citizens are often marginalised. They are not usually the target user group of services based on use of Open Government Data, and even if they are, the type of data opened by public bodies may not be relevant to their needs. Jarke argues that there is a real need to 'bring together city administrations as data owners, technology developers and older citizens as knowledgeable individuals and prospective users in order to co-create relevant and meaningful public services based on open data' (2019: X). In the paper, Jarke presents an evaluation of the innovative "data walks" methodology that she used as a way to engage older citizens in the co-creation of open data based digital services. With particular attention paid to issues of inclusion and efficacy in engaging with the needs and interest of participants, Jarke argues that data walk workshops are one way in which "older adults cease to be subjects of digital innovation and become co-designers" (2019: X). They are a step towards the development of "effective and relevant services for older adults... based on the needs and requirements of the target audience" (p. X16X), rather than being driven by the data that is made available under Open Government Data initiatives.

PART THREE: SHAPING.

The third and final theme to emerge through this special issue is 'shaping'. Three papers in this Special Issue - 'The tower of Babel problem: Making data make sense with Basic Formal Ontology' (Iliadis, 2019), "'Warning! You're entering a sick zone": The construction of risk and privacy implications of disease tracking apps' (Mitchell, 2019), and 'The compliant environment: Conformity, data processing and increasing inequality in UK Higher Education' (Andrews, 2019) – approach their contributions by accounting for the *relationship* between data power and the social. As Baym reminds us, accounts like these locate causality not with technologies themselves or with the people who use them but in the 'middle ground' (2010:44): an approach called the 'social shaping of technology'. The final three contributions to our Special Issue emphasise the intricate interplay between new forms of data (and their systems, management, and so on) and the social world. As Baym notes, perspectives like these tell us that: 'the consequences of technologies arise from a mix of "affordances" – the social capabilities technological qualities enable – and the unexpected and emergent ways that people make use of those affordances' (2010:44).

Andrew Iliadis' research engages with the social shaping and implications of what he calls 'Applied Computational Ontologies' – the 'largely invisible' (p. X4X) standardised metadata vocabularies increasingly used to structure data through organising and labelling, often in an effort to facilitate data integration and interoperability. With a focus on Basic Formal Ontology (BFO) and its various applications, Iliadis adopts a data assemblage approach to illuminate ACOs as "products of human-centred communication" (p. 6). Iliadis' research proposes and puts into practice a methodological approach that combines digital ethnography and digital methods to interrogate critically processes of data labelling and organisation that are under explored across many contexts of application. Examining the ontological realism of such vocabularies, Iliadis argues:

If the ontological realism endorsed by BFO is dependent on the universal laws and structures of science, social ontology is dependent on the invisible rules and laws that society follows...BFO follows social

ontology and has continued the practice by describing institutional systems to which documents belong, positional roles within such systems, and the production of documents...Yet, social kinds are said to be...subjective and dependent on mental attitudes...interactive and malleable. It is here where social ontology gets tricky. (p. X18X)

Through consideration of applications such as the BFO-powered Military Ontology, Iliadis argues that this raises significant concerns about the material impacts of ACOs, concluding with a call for ethical analysis of the BFO methodology when applied in social contexts, and for more research about the interactions between data scientists and ontologists.

Scott Mitchell's article examines the privacy and surveillance implications of digital disease tracking apps, like SickWeather and HealthMap. These apps use data mining, analytics, and crowdsourced data to predict disease outbreaks, and have been called the 'Facebook for hypochondriacs' (Mitchell, 2019: X). The apps work by collecting information from social media and across the wider Web, paired with self-reports from those who use the app to allow users to see who is sick in their neighborhood. HealthMap, for example, was hailed as a 'big data success story' for picking up references to the 2014 Ebola outbreak in Guinea a week before its government notified the World Health Organization (WHO) (Mitchell, 2019). Mitchell explores the discursive dimensions of the apps, noting that disease tracking apps like the above 'construct disease threat as omnipresent and urgent, compelling users to submit personal information – including sensitive health data – with little oversight or regulation' (2019: X). SickWeather, for example, urges its users to help the app to work better – and therefore to raise broader awareness of sickness – by reporting illnesses, which in turn generates data and, eventually, profit for the platform. Influenced by the app walkthrough method (Light et al, 2016), Mitchell found that SickWeather uses smartphone notifications and alerts to create a sense of urgency, discursively compelling users to check the app and add personal information. As Mitchell explains:

SickWeather works to discursively construct disease threat as an omnipresent, inescapable reality, placing contagion fear in users' pocket or the palm of their hand; alerted by a sound from their phone, at any moment their supposed disease risk can change and instantly make itself known. (2019, X).

Mitchell hypothesises that the apps' dominant discourses might shape data disclosure. This paper presents important findings about the implications of a set of under-explored apps (disease trackers) to contribute to broader discussions about social and technical relations.

Penny Andrews' article on the role of data processing in the enforcing of the 'Compliant environment' in UK Higher Education (HE) concludes this Special Issue. The 'Compliant Environment' is a UK Home Office approach to managing immigration, which in part depends upon the reuse and processing of data from a variety of sources. Andrews argues that such a form of 'Compliant environment' is also shaping data practices within the UK's Higher Education system, which is heavily dependent upon data processing for the production of compliance and conformity amongst staff and students. Taking aim at data-driven research metrics, university rankings, student monitoring, evaluation processes, Andrews critically examines the political economy of the data systems that have become common place in HE settings, observing it is no longer the case that 'if you are not paying, you are the product'; rather what we are seeing is 'part-payment via data' (p. X9X) as the business model for many of the data-driven HE systems. Andrews goes on to illuminate the disproportionate impacts of data-driven compliance within HE for marginalised students and staff, including those impacts that result from what can become normalised as a result of HE institutions obligation to monitor students and staff to ensure compliance with immigration authorities. Andrews concludes by calling for resistance to damaging uses of data that 'serve other agendas', advocating a 'civic hygiene' (Schneier, 2007) approach to data management within the HE sector, bringing us back to the question of 'tactics' discussed in earlier papers (Lee and Paris et al).

DATA POWER: DIVERSIFYING THE FIELD.

The papers in this Special Issue address how different forms of online information systems that enable data collection, processing, sharing and use are embedded within, and have implications for, the future development of organisations, cultures and societies. The authors in this special issue approach the topic from a variety of theoretical and methodological angles – some that will be more, and some less, familiar to readers of *Online Information Review*. Our efforts to bring together these papers in a single issues are not only aimed at emphasising a shared focus within the emergent multi-disciplinary critical data studies literature on understanding and resisting data-related injustices, but also to reflect the growing diversity within the field. Among the papers in this issue, readers will find variety in terms of research methods (action research, story completion methods, digital methods, data visualization, interviews, document analysis), geographical reach (Canada, China, Germany, Kosovo, UK, US), and cases (China’s social credit system, activists’ data management systems, disease tracking apps, publicly available Airbnb data, Open Government Data); a diversity that suggests a field that is broadening in scope and depth in an effort to address the pressing societal challenge of widespread datafication.

REFERENCES.

- Baym, N. (2010). *Personal connections in the digital age*. Malden, MA: Polity Press.
- boyd, d. and Crawford, K. (2012). Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication and Society*. 15(5): 662-679.
- de Certeau, M. (1988). *The practice of everyday life*. Berkeley, CA; London: University of California Press.
- Eubanks, V. (2018). *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York: St. Martin’s Press.
- Kennedy, H. (2016). *Post, mine, repeat: social media data mining becomes ordinary*. Basingstoke: Palgrave Macmillan.

Light, B., Burgess, J. and Duguay, S. (2016). The walkthrough method: an approach to the study of apps. *New Media and Society*. 20(3): 881–900.

Lupton, D. and Williamson, B. (2017). The datafied child: the dataveillance of children and implications for their rights. *New Media and Society*. 19(5): 780-794.

Lyon, D. (2014). Surveillance, Snowden, and big data: capacities, consequences, critique. *Big Data and Society*. 1(2): 1-13.

Noble, S. (2018). *Algorithms of oppression: how search engines reinforce racism*. New York: NYU Press.

O’Neil, C. (2016). *Weapons of math destruction: how Big Data increases inequality and threatens democracy*. New York: Crown.

Schneier, B. (2007). Strong laws, smart tech can stop abusive “Data Reuse.” *WIRED*. Available at: <https://www.wired.com/2007/06/securitymatters-0628/> [Accessed May 9, 2018].

Van Dijck, J., Poell, T. and De Waal, M. (2018). *The Platform Society: public values in a connective world*. Oxford: Oxford University Press.