

1 Data and the city

*Rob Kitchin, Tracey P. Lauriault and
Gavin McArdle*

Introduction

There is a long history of governments, businesses, science and citizens producing and utilizing data in order to monitor, regulate, profit from, and make sense of the urban world. Data have traditionally been time-consuming and costly to generate, analyse and interpret, and generally have provided static, often coarse, snapshots of urban phenomena. Recently, however, we have entered the age of big data, with data related to knowing and governing cities increasingly becoming a deluge; a wide, deep torrent of timely, varied, resolute and relational data (Kitchin 2014a; Batty 2016). This has been accompanied by an opening up of state data, and to a much lesser degree, business data, the production of volunteered geographic information, and the emergence of open data cultures and practices (Goodchild 2007; Bates 2012). As a result, evermore aspects of everyday life – work, consumption, travel, communication, leisure – and the worlds we inhabit are being captured and stored as data, made sense of through new data analytics, mediated through data-driven technologies, normalized through data-driven infrastructures, and shared through data infrastructures and data brokers (Amoore 2013; Kitchin 2014b; Offenhuber and Ratti 2014).

This data revolution has produced multiple challenges that require critical and technical attention – how best to produce, manage, analyse and act on urban big and open data, make sense of data infrastructures, data cultures and practices, and understand their consequences with respect to city governance, economy, politics and everyday life. However, to date, there has been relatively little *critical* reflection on the new emerging relationship between data and the city, and how we come to know and understand cities through data in the present era.

In the rush to create so-called ‘smart cities’, wherein core city services and infrastructures become digitally mediated and data-driven – generating, processing and acting on data in real-time to algorithmically manage systems and calibrate performance – much of the attention has been on how to technically create and implement suitable smart city technologies, and associated institutional and infrastructural supports such as data standards, protocols, policies, and a variety of telecom networks. Such data-driven technologies include: urban control rooms, e-government systems, city operating systems, coordinated emergency

response systems, intelligent transport systems, integrated ticketing, real-time passenger information, smart parking, fleet and logistics management, city dashboards, predictive policing, digital surveillance, energy smart grids, smart meters, smart lighting, sensor networks, building management systems, a wide plethora of locative and spatial media. Collectively these technologies are generating an ever-growing tsunami of indexical data (uniquely linked to people, objects, territories, transactions) that can be repurposed in diverse ways – for example, in predictive profiling and social sorting of citizens and neighbourhoods, creating urban models and simulations, for policing and security purposes, etc. (CIPPIC 2006; Batty 2013; Kitchin 2014b; 2016). These data are in addition to large quantities of administrative and statistical data, more traditional sampled survey data, polling and public opinion data, and any other data the city may collect as part of reporting and delivering services.

Rather less attention has been paid to more epistemological, normative, ethical and political questions concerning how data-driven cities and urban issues are framed and approached; how city development and progress are envisaged; what kinds of data are being produced and to what purposes they are being employed; what kinds of cities we ideally want to create and live in (not simply from an instrumental perspective – solving particular issues such as traffic congestion; but with respect to issues such as fairness, equity, justice, citizenship, democracy and governance); how these data-driven technologies and processes work in practice on the ground; what kinds of social and spatial relations they produce; whom they benefit and disadvantage or exclude; what kinds of subjectivity, citizenship, participation and political action they support; and how they reshape many aspects of urban life. This is not to say that there has been no consideration of such questions – as the chapters that follow and the work they reference attest, there is a growing body of research that critically examines urban data and their use. However, the work to date is still relatively formative in theoretical and empirical terms, often considers data-driven systems with the context of smart cities in general terms rather than focusing specifically on the unfolding relationship between data and cities, and the development and rollout of data-driven urbanism is largely outpacing critical reflection and interventions.

Data and the city

This volume is designed to help to fill this lacuna through an interdisciplinary examination of the relationship between data and contemporary urbanism. The focus is not smart city technologies per se, but rather the essays concentrate on how to make sense of urban data and the emerging era of data-driven urbanism. As well as providing synoptic analyses and new conceptual thinking, the chapters detail a number of illustrative examples of urban data, data-driven systems and related issues, including data infrastructures, urban blockchains, mapping, urban modelling, data provenance, data quality, data citizenship, citizen science, data practices, data cultures, data frictions and city dashboards. Importantly, given the wide-ranging, diverse and complex relationship between data and the city, and

the need to bring various expertise and knowledge into dialogue, the contributors are drawn from a number of disciplines (Geography, Geographic Information Science, Planning, Sociology, Information Science, Design, Media Studies, Law and Computer Science).

All but three of the chapters were prepared initially for a workshop at the National University of Ireland Maynooth in September 2015, funded by the European Research Council through an Advanced Investigator Award to Rob Kitchin for The Programmable City project (ERC-2012-AdG-323636-SOFTCITY). Each essay was pre-prepared and submitted in advance of the meeting, then extensively discussed at the workshop, and subsequently revised for publication. While the book is designed to work as a standalone text, there is a companion book, *Code and the City* (Kitchin and Peng 2016), that focuses predominately on the relationship between software and the city. To provide a structure, we have divided the book into four parts.

Data-driven cities

The first part considers the relationship between data and the city in a broad sense, focusing on the creation of real-time cities and data-driven urbanism and how the ever-greater flows of data are transforming city services, infrastructures, urban life and how we understand and govern cities.

In the opening chapter, Martijn de Waal examines the creation of ‘real-time cities’, wherein computation is embedded into the fabric of cities producing real-time data flows that can be used to know and manage city services in the here-and-now. He argues that such data-driven systems are changing how we understand cities in three ways. The first is the adoption of an action-orientated epistemology wherein the production of real-time data, along with machine learning techniques, enables a new kind of scientific knowledge about cities that treats them as complex systems which can be made actionable through smart city technologies. The second approach is more critical in orientation and, on the one hand, challenges the scientific principles and epistemology of the first, and on the other, considers more ontological questions concerning how real-time data and data-driven systems transform the production of space, the nature of place, and the experience of living in the city. The third approach asks more normative questions and argues that cities cannot be conceptualized and approached as being analogous to other complex systems, such as galaxies and rainforests, because they are social-cultural-political in nature. Instead, it is contended that a new science of cities needs to frame data-driven cities with respect to wider concerns about the kinds of cities we want to create and how to produce particular kinds of ‘cityness’. De Waal argues that more attention needs to be paid to this third kind of knowledge making and its praxes.

Mike Batty considers the nature of urban big data and the epistemological challenges of using them to make sense of the city, placing his discussion in historical context. Adopting an approach that is perhaps characterised as fitting within de Waal’s first mode of understanding data-driven cities, Batty argues that we have

always been struggling to extract insights from ever-larger and more dynamic data as urban technologies evolve and urban computational research struggles to keep up. He notes that what might be considered small data – sampled in time, space and by category – soon become very large once the interactions between data points are examined. Using the concept of a data cube, Batty examines the characteristics of urban flow data between locations. In particular, he illustrates his arguments by detailing the difficulties of making sense of traditional transport interaction data, such as origin (home) to destination (work) flows across a city, and more dynamic and massive datasets, such as the tap-in and tap-outs of travelers on the London Underground (one of his datasets consists of nearly 10 billion records generated over 86 days in the summer of 2012). In both cases, urban science is still struggling to extract and communicate meaningful insight. He concludes that rather than abandoning theory for an empiricist form of data science, there is a pressing need to develop a theoretically insightful urban science.

In his chapter, Rob Kitchin argues that while there has long been forms of urbanism that are data-informed, a new era of data-driven urbanism unfolding as cities become ever more instrumented and networked, their systems inter-linked and integrated, and vast troves of big urban data are being generated and used to manage and control urban life in real-time. He contends that data-driven urbanism is the key mode of production for what have widely been termed smart cities. Adopting an approach that largely maps onto de Waal's third approach, Kitchin critically examines a number of urban data issues, including: the politics of urban data and production of data assemblages; data ownership, data control, data coverage and access; the creation of buggy, brittle, hackable urban systems (data security, data integrity); and social, political, ethical effects (data protection and privacy, dataveillance, and data uses including social sorting and anticipatory governance). He concludes that whilst data-driven urbanism purports to produce a common-sense, pragmatic, neutral, apolitical, evidence-based form of responsive urban governance, it is nonetheless selective, crafted, flawed, normative and politically inflected. Consequently, whilst data-driven urbanism provides a set of solutions for urban problems, it does so within limitations and in the service of particular interests or there is an overreliance on mathematically and engineered models that do not factor in a city's social, cultural, historical, institutional and political complexities. Those very things that give cities their character.

Urban data

The second part focuses attention on the nature of urban data, examining them from ontological, political, practical and technical points of view. Importantly, the analysis does not conceive of urban data from a common-sense, essentialist position, wherein they are seen to faithfully and validly represent the state of the world, but rather consider the ways in which data are produced and framed within socio-technical systems.

Teresa Scassa provides a critical overview of crime data and their sharing through open data sites, interactive visualizations, and other media. She details

how crime data are far from neutral, objective records of criminal, policing and legal activity, but rather are shaped significantly by legal, institutional and cultural factors. She argues that crime data are subjective and contested, record certain kinds of information but excludes others, and are known to be full of gaps and errors. Moreover, capturing, analysing and acting upon crime data requires human interpretation and judgement, framed with societal and institutional contexts. And yet, despite these issues, crime data are often taken at face value and are used to drive social, policing, security and legal policy and programmes and to underpin new interventions such as predictive policing. While the data do hold value and are important in revealing levels of crime and society's institutional response, she contends that they need to be treated with caution, with users considering how, by whom, and for what purposes the data were generated to gauge their veracity and trustworthiness.

Jim Thatcher and Craig Dalton similarly consider the issues of data veracity and trustworthiness by considering data provenance. They note that data provenance is presently largely instrumental in nature and concerns information about the production and history of a dataset. Such information allows users to know how the data were captured, by whom, using what techniques and technologies, how they were processed and handled, and so on, enabling them to judge their quality, shortcomings and suitability for use. Typically, such information is stored as a metadata – that is, data about the data. However, they contend that such an instrumental approach to data provenance is limited and too technically orientated, ignoring the wider context in which the data are produced and used. Instead, they suggest the use of a more-than-technical form of provenance that not only documents traditional metadata, but also includes situated contextual factors such as motivation, value and power. They formulate this version of data provenance as the recording of 'data encounters' which capture the always already-cooked nature of data and the contextual nature of its use.

Jim Merricks White likewise is interested in data encounters, but rather than focus on provenance, he seeks to follow data from their generation through to their various uses, exposing how they are cleaned, recombined and put to work. Using an empirical example of infant mortality and their use in city indicator initiatives he charts the translation and circulation of data, seeking to document what he terms 'data threads', highlighting the entanglement of data infrastructures and geography, and their inherent materiality and relationality. He traces how infant mortality data are generated by messy human and computational practices shaped by a framework of definitions and standards. These data are then used in varying ways, reworked to create new derived data, and used in ways not anticipated with respect to their original generation. He notes that the devastating loss of a child's life is rendered first as trace, then as data point, and then as input to derivative calculations and distant ambitions, in this case various health and city indicator initiatives. With each transformation, he argues the data become increasingly alienated from their material associations and their meaning mutate to reflect new discourses and ideologies. Comparing his notion of data threads to that of 'data journeys' detailed by Bates *et al.* (2016), White provides a useful epistemological

avenue for thickening the description of data assemblages and how data translate and are woven together across such assemblages.

Considering the nature of urban data further, Dietmar Offenhuber examines what makes urban data meaningful, the extent to which data are always cooked and never raw, and concerns with respect to the repurposing data. Utilizing the concept of ‘data friction’ he examines the issues that arise when data and metadata generated by different organizations, that utilize different formats and standards, are moved or bought into contact. He notes that despite difficulties and limitations, data sets can develop a life of their own and be repurposed in diverse ways, often as data proxies for other phenomena. Offenhuber examines these issues with respect to Twitter data, which have become widely used in social science research, and satellite imagery generated by the Operational Linescan System (OLS) of the US Air Force’s Defense Meteorological Satellite Program (DMSP). He contends that Twitter data, despite its widespread repurposing, are ‘sticky data’, that is meaningful when discussed in their original context, but problematic to interpret, extrapolate and generalize otherwise. In contrast, OLS/DMSP data are relatively non-sticky, being used extensively to identify city street lighting and act as a proxy for population density and economic activity, though it is not without problems. Offenhuber thus concludes that as proxies for urban phenomena, both data sources offer only partial perspectives and need to be used with caution.

Urban data technologies and infrastructures

The third part examines the constellation of existing and emerging urban data technologies and infrastructures. The chapters explore a range of political, practical and technical issues and epistemological and theoretical approaches with respect to building, operating and making sense of such data-driven systems.

One way in which a plethora of urban data are made sense of by city managers and shared with citizens is through city dashboards that provide a variety of visualization and analytic tools which enable these data to be explored. While such dashboards provide useful tools for evaluating and managing urban services, understanding and formulating policy, and creating public knowledge and counter-narratives, Rob Kitchin and Gavin McArdle’s analysis reveals a number of conceptual and practical shortcomings. They critically examine six issues with respect to the building and use of city dashboards: epistemology, scope and access, veracity and validity, usability and literacy, use and utility, and ethics. Drawing on their experience of building the Dublin Dashboard, they advocate a shift in thinking and praxis that openly situates the epistemology and instrumental rationality of city dashboards and addresses more technical shortcomings.

Pouria Amirian and Anahid Basiri also consider the sharing and analysis of urban big data, though their focus is more technical in nature. Given the wide variety of different data-driven platforms being utilized across a number of organizations and domains, and the need to be able to share and integrate such data so they can be used by many systems and actors, it is necessary to create platform-independent principles and mechanisms to ensure interoperability. They contend

that such interoperability is best achieved through Service Orientation Principles (SOP) along with a new architecture, Organizational Service Layer, that uses polyglot binding. They detail three core SOP approaches, and their benefits and shortcomings, currently being utilized to share data and analysis (Web Services, RESTful services and Geoservices), as well detailing how four types of bindings can be used to provide loose couplings between backend implementation and other software applications. These bindings enable platform independency and agile and straightforward communication between systems, thus creating accessible, flexible, scalable and interoperable smart city platforms and more easily implementable city data portals, urban control rooms and city dashboards.

An alternative and emerging form of data infrastructure for city dashboards and services are blockchains. Blockchains are sealed and encrypted distributed ledgers of all transactions ever conducted within a system. Each block records key metadata regarding a transaction such as information about sender and receiver, time, value, fees and IP address, and once recorded cannot be altered, thus creating trust. Each block adds to the sequence of transactions forming a chain that leads back to the start of the database. While blockchains are mostly commonly associated with new financial currencies such as Bitcoin, Chris Speed, Deborah Maxwell and Larissa Pschetz examine their utility for recording and sharing other kinds of transactions. To illustrate how blockchains work as economic, social and cognitive ledgers they discuss their use with regards to finance and work. They then detail the development of two prototype city ledgers produced in a design workshop that utilize Bitcoin technology demonstrating how blockchains offer opportunities to capture diverse social practices and transactions in city ledgers. They contend that the blockchain has the potential to create trusted city ledgers (databases), and thus trusted city dashboards, and provide the foundation for dealing with complexity and predicting future outcomes.

Rather than focus on the form, operation, building and shortcomings of building data infrastructures, Till Straube focuses on how best to theoretically and empirically make sense of them. He proposes a materialist approach to understanding the constitution and work of data infrastructures and data-driven systems. Instead of concentrating on the relational effects of such infrastructures – how they produce space–time compression or a space of flows – he argues that attention needs to be paid to the materiality and spatiality of the infrastructures themselves (programming languages, database software, data formats, protocols, APIs, etc.). Such a focus, he argues, foregrounds data technologies and infrastructures, their make-up and practices, and how they are materially embedded into the fabric of cities and everyday lives. The approach he advocates is a topological reading of data technologies, underpinned by assemblage theory. Here, emphasis is put on charting the network of relations between potentially dispersed socio-technical systems, rather than the topography of their territories; that is, it is concerned with material connections and power relations that operate across and produce a relational rather than Euclidean space. He thus forwards an epistemology, what he terms an applied materialist topology, that seeks to pay close attention to how data technologies and infrastructures articulate, perform and translate time-spaces

within a socio-political context. Such an approach also makes clear that as well as having a materiality, data technologies are never neutral in formulation, operation and effects.

Tracey Lauriault also aims to make sense of data infrastructures and offers a nested methodological approach to study the power/knowledge of data models and ontologies. Drawing on ethnographic work in which she was embedded at Ordnance Survey Ireland (OSi) examining how the organization introduced a fundamentally new data model, Prime2, that replaced a map layers model with an object-orientated model, she considers how cities are captured within data models and how these models transduce the city. She advances three interlinked methodological approaches for making sense of the diverse range of empirical materials she amassed, including interviews, technical documents, procedure and training manuals, databases, in-field observation and news reports. The first is the application of her modified version of Hacking's dynamic nominalism to assess how the city is 'made-up' through the new Prime2 ontology. The second is a genealogical mapping of the development of the Prime2 data model over time and the key events in its production. The third is an application of Kitchin's socio-technical assemblage as a framing tool to study how the model constitutes one part of a national spatial data infrastructure. She argues that using these methodological approaches together enables an unpacking of the discursive and material production of data models and data infrastructures and how these models and infrastructures produce space.

Urban data cultures and power

The fourth part considers the social and political configurations of urban data infrastructures and data-driven systems and who they are operated by, their purposes and who they serve. Far from being neutral and objective in nature and serving the public good in a general sense, this part examines their data cultures and data power.

As Jo Bates notes, data do not arise from nowhere. Rather, data are produced by people and technology embedded within socio-material relations situated within time and space. They are the result of data practices and modes of data governance operating within specific data cultures. In other words, data production and use is shaped by cultural norms, value systems and beliefs, as well as the wider political economy and institutional and legal landscape. Data cultures, and their sites of practice and governance, are historically constituted, dynamic, open and porous, and thus mutate over time. Bates notes that for each city there are a multitude of interrelated data cultures operating within and across public organizations, private enterprise and civic bodies, though these cultures are not all created equal, with some dominating and subverting others. She argues that it is important to unpack these data cultures and their sites of data practice to reveal their assumptions, values, participants, rhetorical and material work, the power dynamics at play, how they shape the domain on which they operate, and how they interconnect with other data cultures. In so doing, the inherent politics and

power of such systems are revealed, enabling us to challenge and reconsider how they are conceived and work in practice.

Given that data about the city and its citizens are produced to enable the functioning of city systems, monitor and regulate populations, to underpin markets, or to provide counter-systems they are fundamentally instruments of power and capital (even when they seemingly enable diverse communication, communities and play – there are always inclusions and exclusions in their production and whom benefits from their operation). They therefore raise important questions concerning citizenship and political subjectivity in the digital age. Evelyn Ruppert examines the extent to which people are data subjects or data citizens in the contemporary era and how data citizenship is constituted. She argues that to understand the relationship between data and the city necessitates asking political questions concerning the framing, identity and positioning of digital subjects and the conduits of power that systems work to reproduce. The data of cities, she notes, are produced by technologies in the employ of public institutions and companies that confer differing forms of citizenship, though these are not accepted uncontested. This is evident in ongoing debates concerning the production of big data and surveillance, privacy, confidentiality, anonymity, security, policing, governance and data markets. Rather than focusing on the substantive nature of digital data rights, Ruppert concentrates on who are the subjects of these rights, their political subjectivity, and the role of subjects in the making and shaping of data, developing a theory of data citizens.

Muki Haklay approaches the question of citizenship through the emergence of citizen science. His starting questions are to ask: whether the future being produced within the smart city vision by data-driven technologies is the one citizens want? And whether such technologies integrate and foster meaningful and purposeful social and communal activities or create feelings of alienation? His concern is that smart city systems represent the interests of city governments and corporate interests and focus on instrumental issues rather than human values and desires; on technocratic constraints and management rather than imagination and serendipity. He explores these issues drawing on the ideas of Albert Borgmann, especially those relating to the difference between device paradigms (instrumental, technically mediated engagements) and focal practices (meaningful social engagement). He argues that smart city technologies tend towards the former, being automated and autocratic, whereas as citizen science initiatives tend towards the latter, being more social and community engaged. He thus argues for a more open, democratic and participatory vision of data-driven city systems in which people play an active role as citizens, not simply subjects. Such meaningful participatory and collective action centred on focal practices, he contends, has the potential to transform the present smart city paradigm.

Future agendas

Taken together, the chapters highlight the diverse ways in which data and cities are becoming ever more intertwined, transforming how we come to know, manage,

govern and live in cities. There are several themes that cut across the essays and in conclusion we want to highlight three that we believe require particular theoretic and empirical attention.

Data politics and power

Collectively, the chapters that follow make a compelling argument that urban data are always cooked and never raw, and the data-driven systems and infrastructures that produce, manage, share and act on them are socio-technical systems not simply technical ones. Urban data and systems then are never neutral, objective and common-sensical, but rather are inherently political – invested with values and judgements, are formed and operated within cultural milieu, and are designed to produce certain effects. This is as much the case for initiatives that seek to be inclusive and enable citizen-engaged data projects, such as open data sites, as it is for systems designed for state surveillance or corporate profit. Certainly, many data-driven city systems and their data practices work to manage, regulate and control urban activities; they inherently capture certain kinds of data and use them to enact particular power relations. Much rhetorical and material work is invested in reproducing the logic and legitimacy of these systems, for example through smart city discourses, but they always remain open to resistance, subversion and transgression.

Data and the City performs important work, we believe, in exploring urban data politics, cultures and power. However, there is still much empirical and theoretical research needed to unpack the specific ways in which data are cooked and utilized to perform political work, however subtle that may be – to examine: how data are generated, processed, shared, translated and used; how data cultures form and are reproduced; how data practices operate within and across networks of actors and data-driven technologies; how data-driven systems produce political subjectivity and data citizens; how data cultures, politics and practices create ethical dilemmas, especially with regards to dataveillance and the work of data brokers; and the forms and practices of alternative data-driven systems that seek to enact more participatory and emancipatory politics. Moreover, further research is required to understand how data influence digital labour, investigating issues such as how institutional and organizational structures change with the introduction of new databased regimes, how data ecosystems change government and corporate work practices, and how the database managers and data scientists become more important within institutions with their knowledge and expertise becoming privileged over others.

Epistemology

As the chapters make clear, there are a diverse set of epistemologies being deployed to make sense of urban data, data-driven systems, and the relationship between data and the city. This varies from more computational approaches, such as the urban science practised by Batty and the technical mapping of data-driven systems

by Amirian and Basiri, through to the Kitchin's unpacking of data assemblages, Straube's applied materialist topology, Lauriault's modified use of dynamic nominalism and genealogical approach, Merrick White's strategy of following data threads, Bates's mapping of data cultures, and Haklay's charting of citizen science. These are by no means the only epistemological approaches being used, as illustrated by de Waal's chapter. The sheer variety of disciplinary and philosophical traditions, technologies and issues make this epistemological diversity inevitable, and we would not be in favour of trying to advocate for a single epistemological paradigm. We do, however, believe that much more attention needs to be devoted to the epistemological challenge of providing useful insights into urban data systems and infrastructure and data-driven urbanism.

These challenges include trying to make sense of highly dynamic, complex and capricious domains that are full of various actors and actants, interlinked systems, diverse practices and processes, and competing politics and interests, and are often black-boxed (in terms of the technical processes, but also institutional access). Moreover, these domains work across scales from single devices to entire cities. Indeed, there is a need for a balance between detailed and empirically rich mappings of individual systems that tease apart their complex relations and workings, and wider synoptic analysis of how these data-driven systems and cultures are working together or in conflict to produce data-driven urbanism. The pressing task then is to, on the one hand, develop conceptual tools for making sense of data-driven technologies and urbanism, their architecture and workings, and the transformations they are producing, and on the other to identify suitable methodologies for grounding such tools through empirical research. While the chapters provide some useful starting points, building on longer legacies of related research, it is clear that there is much more to be done.

Normative questions

Most analysis of urban data and data-driven city systems grapple little with normative questions – that is, consider in-depth questions about how things should be as opposed to how they are, or for whom and what purpose are data-driven cities being created? Instead, analysis is concerned with detailing how systems are configured and work, either from a technical or social perspective. For those that develop such systems, the goals are usually defined in instrumental terms – to make a city more sustainable, resilient, efficient, secure, competitive, and so on. As a consequence, a fundamental question such as 'what kind of cities do we want to create and live in?' has largely been framed technically and instrumentally, rather philosophically in relation to issues such as fairness, equity, justice, citizenship and democracy. By highlighting issues such as data cultures, data power and data citizens, the chapters in *Data and the City* point to the need to consider wider normative questions about the goals of data-driven urbanism and whose interests they should serve. For example, should data-driven systems be primarily about creating new markets and profit? Facilitating state control and regulation? Improving the quality of life of citizens? Or all

three and in what balance? And in what form should they be conceived and implemented? Exploring, debating and answering these normative questions is important because they frame how the urban data revolution will unfold and how policy and law will need to be formulated to produce the kinds of cities desired. In fact, framing the debate in instrumental terms has been a useful rhetorical strategy for avoiding such normative debate because it shifts the debate into a post-political and seemingly common-sensical register. We believe it is time to challenge such a positioning.

Conclusion

As we have noted above, there are many political, ethical, epistemological and normative questions still to be asked and answered with respect to urban data, data-driven city systems and urbanism, yet the urban data revolution continues to unfold at pace. There is thus a pressing need for empirical research and conceptual thought to make sense of the changes taking place. Collectively, we believe the chapters in *Data and the City* provide a productive set of routes into thinking about these questions that help advance our understanding of the evolving relationship between data and urban life and forms of data-driven urbanism. As such, it adds to an emerging interdisciplinary body of work and should hopefully make for an illuminating and stimulating read.

Acknowledgements

The research for this chapter and the Data and the City workshop were funded by a European Research Council Advanced Investigator Award to Rob Kitchin, entitled 'The Programmable City' (ERC-2012-AdG-323636-SOFTCITY).

References

- Amoore, L. (2013) *The Politics of Possibility: Risk and Security Beyond Probability*. Durham, NC: Duke University Press.
- Bates, J. (2012) 'This is what modern deregulation looks like: Co-optation and contestation in the shaping of the UK's Open Government Data Initiative', *The Journal of Community Informatics* 8(2), available from: www.ci-journal.net/index.php/ciej/article/view/845/916 [accessed 6 February 2017].
- Bates, J., Lin, Y-W. and Goodale, P. (2016) 'Data journeys: Capturing the socio-material constitution of data objects and flows', *Big Data & Society* 3(2): 1–12.
- Batty, M. (2013) *The New Science of Cities*. Cambridge, MA: MIT Press.
- Batty, M. (ed.) (2016) Big data and the city. Special issue of *Built Environment* 42(3), available from: www.alexandrinepress.co.uk/built-environment/big-data-and-city.
- CIPPIC (2006) *On the Data Trail: How detailed information about you gets into the hands of organizations with whom you have no relationship. A Report on the Canadian Data Brokerage Industry*. The Canadian Internet Policy and Public Interest Clinic, Ottawa. <https://cippic.ca/sites/default/files/May1-06/DatabrokerReport.pdf> [accessed 6 February 2017].

- Goodchild, M.F. (2007) 'Citizens as sensors: the world of volunteered geography', *GeoJournal* 69: 211–221.
- Kitchin, R. (2014a) 'The real-time city? Big data and smart urbanism', *GeoJournal* 79(1): 1–14.
- Kitchin, R. (2014b) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: SAGE.
- Kitchin, R. (2016) *Getting smarter about smart cities: Improving data privacy and data security*. Data Protection Unit, Department of the Taoiseach, Dublin, Ireland, available from: www.taoiseach.gov.ie/eng/Publications/Publications_2016/Smart_Cities_Report_January_2016.pdf.
- Kitchin, R. and Perng, S.Y. (eds) (2016) *Code and the City*. New York: Routledge.
- Offenhuber, D. and Ratti, C. (eds) (2014) *Decoding the City: Urbanism in the Age of Big Data*. Basel: Birkhauser.

Taylor and Francis
Not for distribution