4

Urban science

Prospect and critique

Rob Kitchin

Introducing urban science

Urban science practices and promotes an interdisciplinary scientific and computational approach to city systems and the processes of urbanization. It uses statistical analysis and data analytics – including machine learning, data mining, visual analytics, modelling, and simulation – to identify causal relationships and predict how city systems work. This is in contrast to urban studies, which uses both quantitative and qualitative methods and adopts a more contextual approach with respect to politics, culture, policy, and history. While urban studies generally conceives of cities as constellations of places, with analysis usually based upon fairly static empirical data (small samples, generated at specific places and times), urban science views cities as systems (or a system of systems) with analysis utilizing urban big data (massive samples generated on a continuous basis). Typically, urban science seeks to map and model urban *dynamics* – patterns of flow, urban processes, and system interactions. The aim is to determine urban 'laws', conduct real-time analysis of systems, produce new theoretical insights, develop a synoptic and integrative science of cities, and to translate the knowledge produced into practical application, including urban design and planning, city management, and economic development.

Urban science builds on a longer history of quantitative social science, including quantitative geography, geographic information science, urban and transportation modelling, social physics, urban and regional economics, urban cybernetics, social ecology, and location theory, that have sought to explain and model urban processes and the functioning of city systems (Batty 2013a). However, many academics and industry analysts now practising and promoting urban science have little grounding in this history, with their training being rooted in the fields of data and information science, computer science, physics, and engineering. Moreover, they have little knowledge of the deep history of research in urban studies and allied disciplines of sociology, geography, anthropology, economics, history, architecture, and planning. Urban science researchers have been attracted to investigating urban processes by the massive volumes of urban big data now being generated, the call for science to tackle global challenges such as rapid urbanization, sustainability, and climate action, along with associated research funding streams, and the wider promotion by industry and governments

9781138036673C04.3D 43 [42–50] 14.12.2019 5:20PM

for the creation of 'smart cities' (Townsend 2013, 2015; O'Sullivan and Manson 2015). These drivers have also led to the creation of a number of large, interdisciplinary urban science research centres across the globe (see Batty 2013b; Townsend 2015). Industry and government are often external partners or stakeholders in these centres and their projects, providing data, funding, and other in-kind contributions, and are often direct beneficiaries of the research in terms of new knowledge, intellectual property, products, and networks.

For many of those practising urban science, the approach is promoted as a paradigm-shifting endeavour – where urban science promises to provide a more integrative and insightful understanding of cities than urban studies that will transform how urban policy making and planning is undertaken, and will become the dominant approach for urban research. Indeed, Solecki et al. (2013) argue that urban studies, rooted in a social sciences rather than computational, scientific tradition, has failed to deliver knowledge that effectively solves city issues and is inappropriate for delivering solutions for the major challenges ahead as urbanization continues apace. This unsuitability is due to its disciplinary fragmentation, a panoply of approaches that create disparate viewpoints, and its focus on cities as places and on the symptoms of urban problems. Instead, Solecki et al. (2013) call for an urban science that focuses on urban processes and systems, and underlying causes and potential solutions (not place and symptoms), and shares a common epistemological approach underpinned by a scientific method. They propose three basic goals for urban science:

(1) To detail the basic components of urbanization across scales;
(2) To identify the universal laws of city-building;
(3) To find relationships between urbanization and other aspects of Earth's systems.

Only urban science, Solecki et al. (2013) contend, can produce 'a theory of urbanization with fundamental and unique components that can withstand scientific scrutiny' and 'lead to systemic solutions that address the whole rather than separate components' (p. 14)

Further, because urban science conducts analyses and builds models based on urban big data, it is posited that it offers the potential for urban knowledge that has greater breadth, depth, scale, and timeliness, and is inherently longitudinal, in contrast to that derived from urban studies (Batty *et al.* 2012). Big data have fundamentally different properties to traditional 'small' datasets, being generated in real-time, exhaustive in scope, and having fine resolution (Kitchin 2014). For example, rather than data being derived from a travel survey with a handful of city dwellers during a specific time period at particular locations (i.e., sampled 'small data'), transport big data consist of a continual survey of every traveller, for example, collecting *all* the tap-ins and tap-outs of travel cards, or using automatic number plate recognition-enabled cameras to track *all* vehicles, or using sensors to monitor the mobile phone MAC (media access control) addresses to track *all* pedestrians with a phone. It thus becomes possible to determine detailed patterns of travel across times of the day, days of the week, and seasons, and to do this for all nodes on the network (e.g., junctions, bus stops, sensor locations), and to make predictions about future system performance under different conditions. As a consequence, data from such systems have the potential to produce a highly granular, longitudinal, whole system understanding of a city system and enable it to be managed in real time.

Over the past couple of decades, this transformation from slow and sampled 'small' data to fast and exhaustive 'big' data has been enabled by the roll-out of a raft of new networked, digital technologies embedded and integrated into the fabric of urban environments and infrastructures. Such technologies include digital cameras, sensors, transponders, meters, actuators, and GPS that monitor various phenomena and continually send data to an array of control and management systems, such as city operating systems, centralized control rooms, intelligent transport systems, logistics management systems, smart energy grids, and building management systems. In addition, a multitude of smartphone apps and sharing economy platforms generate a range of real-time location, movement, and activity data. The result is a deluge of real-time, fine-grained, contextual, and actionable data which are routinely generated about cities and their citizens upon which urban science can be practised (Koonin and Holland 2013). However, such data are largely closed in terms of access given that they are mostly generated by privately owned systems, meaning that practising urban science often requires developing access rights with companies and states.

The remainder of the chapter sets out urban science's relationship to urban informatics and explains its epistemology. It then summarizes criticism of urban science with respect to epistemology, instrumental rationality, data issues, and ethics.

Urban science and its relationship to urban informatics

There seems to be some confusion in the literature as to the relationship between urban science and urban informatics, a term that pre-dates urban science and is used to describe a form of academic enterprise concerning the generation, management, processing, analysis, and utilization of urban data (Foth 2009). It is worth teasing out the overlap and differences here in order to make clear the nature of urban science. While Batty (2013b) frames urban science within a larger domain of urban informatics, Townsend (2015) positions urban informatics as sub-branch of urban science, and Kitchin (2016) has them as separate but complementary fields that often intersect. This confusion is due to how urban informatics has been conceived.

For Batty (2013b: 3) urban informatics is the 'application of computers to the functioning of cities' and 'the ways in which computers are being embedded into cities'. Here, urban science is one way, within the broader remit of urban informatics, in which computation is being utilized to understand the functioning of cities and in turn informs how computation is used to manage and control urban systems. For Foth (2009) urban informatics is an interdisciplinary enterprise that includes three broad communities: social (e.g., media studies, communication studies, cultural studies, sociology); urban (e.g., urban studies, geography, planning, architecture), and the technical (e.g., computer science, data science, electronic engineering, human–computer interaction). From this perspective, urban informatics is primarily concerned with the development of informational tools and management systems for controlling and communicating urban processes, understanding human interactions with such systems, and studying the relationship between people, place, and digital technology (Foth 2009), rather than being centrally concerned with urban modelling, statistical analysis, simulation and prediction, and finding 'urban laws'. Moreover, urban informatics given its wider body of constituent practitioners is less likely to be positivistic in nature or follow the 'scientific method'.

Many of the new urban research centres detailed by Batty (2013b) and Townsend (2015) conduct both urban science and urban informatics research. For example, the Centre for Advanced Spatial Analysis (CASA) in University College London undertakes a range of applied and fundamental geospatial research focused on modelling and simulating cities, including creating 3D virtual models and city dashboards. The Centre for Urban Science and Progress (CUSP) at New York University seeks to use a scientific approach to develop data-driven solutions for explaining and tackling urban problems and offers a Masters programme in Urban Science and Informatics.

Urban science and its epistemologies

Urban science is broadly rooted in a positivistic tradition that has sought to apply scientific principles and methods, drawn from the natural, hard, and computing sciences, to social phenomena in order to explain them. The aim is to statistically test relationships between variables or build models as a means to produce and verify laws that explain and predict how systems work. Central to this endeavour is the objective collection of data through common and standardized methods of observation (that can be replicated) and the formulation of theories which can be tested and verified. In general, a realist epistemology is adopted that supposes the existence of an external reality which operates independently of an observer and which can be objectively and accurately measured, tracked, statistically analysed, modelled, and visualized to reveal the world as it actually is (Kitchin *et al.* 2015). In other words, it is held that urban data can be abstracted from the world in neutral, value-free, and objective ways and are understood to be essential in nature. That is, data are representative of that which is being measured, faithfully capturing its essence and are independent of the measuring process (though it is acknowledged that there might be data quality issues related to error, bias, calibration, etc.). These data, when analysed in similarly objective ways through statistical analysis, modelling, and simulation, reveal deep insights about cities that can be used to reshape urban policy and enhance urban infrastructures (though it is appreciated that there might be constraints and limitations due to the methodology employed). While cybernetic modelling approaches recognize the complexity and emergent qualities of city systems, such systems are still understood in machinic terms and are largely closed and bounded in nature.

Three epistemological variations of urban science are practised (Kitchin 2014). The first is a traditional, hypothesis-driven, deductive scientific method, with questions and approach guided by established theory. The second is a form of inductive empiricism in which it is argued that, by employing data analytics, urban big data can speak for themselves free of theory or human bias or framing (Kitchin 2014, 2016). Such an approach is best exemplified by Anderson (2008: n.p.) who argues that 'the data deluge makes the scientific method obsolete' and that '[c]orrelation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all'. In other words, rather than being guided by theory, the data can be wrangled through hundreds of algorithms to discover the most salient factors with regards to a particular phenomenon. Such an approach has gained some traction in data science and within industry research. The third is data-driven science that seeks to hold to the tenets of the scientific method, but generate hypotheses and insights 'born from the data' rather than 'born from the theory' (Kelling *et al.* 2009). It uses guided knowledge discovery techniques to mine data to identify potential hypotheses, before a traditional deductive approach is employed to test their validity. This approach is more common because it rejects the idea of the 'end of theory' and maintains scientific values; extracts additional, valuable insights that traditional knowledge-driven science would fail to generate; and it produces more holistic and extensive models and theories of entire complex systems rather than elements of them (Kelling *et al.* 2009; Miller 2010).

In many cases, these approaches have been realized through applied research that uses city environments as 'living laboratories'; that is, as testbeds to validate the science and test the practical interventions produced (see Evans *et al.* 2016). For example, CASA uses London as its laboratory and CUSP uses New York, working with public and private stakeholders to tackle the real-world problems they have identified. Indeed, much urban science research is highly empirically grounded and applied in nature, with extensive collaboration between scientists, city administrations/state agencies, and industry partners. The potential benefits to each party are clear – academic gain access to key datasets, companies gain access to intellectual insight, and city administrations/state agencies gain access to potential interventions and solutions. The Smart Dublin initiative in Ireland, for example, offers university researchers and companies (often working in partnership) the opportunity to work with four local authorities and within a couple of designated urban testbed areas to experiment with new technologies, generate urban big data, and practice urban science. At the start of 2019 there were 109 active projects – not all of which were urban science orientated – aimed at understanding and solving a diverse set of urban problems.

Criticism of urban science

While urban science has expanded rapidly in the last decade, it is far from producing a paradigmatic shift in urban research and has been subject to critique from urban studies scholars and others (Crampton *et al.* 2013; Kitchin 2016; Mattern 2013). This critique is multi-pronged, with much of it mirroring that of positivistic social sciences and geographic information sciences in previous decades (see Crampton 2010; Kitchin 2015). In these earlier 'theory wars' urban and spatial science were roundly criticized for being too reductionist, mechanistic, essentialist, and deterministic, collapsing diverse individuals and complex, multi-dimensional social structures and relationships to abstract data points and universal formulae and laws. Moreover, rather than being epistemologically objective, neutral, and value-free, it was demonstrated that such science was framed and situated within power-geometries of knowledge and practice and often served particular interests (Pickles 1994). In addition, they also wilfully ignored the metaphysical aspects of human life and the role of politics, ideology, social structures, capital, and culture in shaping urban relations, governance, and development (Kitchin 2016).

Consequently, scientific approaches to cities have been critiqued as being rather naïve and narrow in perspective, producing overly simplified explanations and models, and a limited and limiting understanding of how cities work (foreclosing what kinds of questions can be asked and how they can be answered) and how urban issues can be tackled. They promote an instrumental rationality that posits that cities can be effectively steered and managed through scientific insights and technical instruments and that urban issues can be solved through a range of technical solutions (Kitchin *et al.* 2015; Mattern 2013). Urban science, it is argued, has thus far failed to recognize that cities are complex, multi-faceted, contingent, relational systems, full of contestation and wicked problems that are not easily captured or steered, and that urban issues are often best solved through political, civic society, fiscal, policy, and legal interventions rather than technical fixes and technocratic forms of governance (Kitchin *et al.* 2015). Indeed, critique of the first wave of cybernetic approaches to cities in the later 1960s and 1970s demonstrated that they produced knowledge and policy interventions that not only failed to live up to their promises but did much damage to city operations (Flood 2011; Townsend 2013). For example, New York's adoption of the RAND Corporation's cybernetic model for the redistribution of fire stations contributed to the destruction caused by fires that blighted the city in the 1970s (Flood 2011).

While advocates of computational social and urban science counter that the availability of big data and data analytics address some of the criticisms of earlier forms – especially those of reductionism and universalism by providing more finely grained, sensitive, and nuanced analysis that can take account of context and contingency (Kitchin 2014) – many concerns undoubtedly still hold for present forms of urban science (Kitchin 2016; Wyly 2014). For example, Batty (2013a, 2013b) notes that, despite drawing on complexity theory and advances in data analytics, urban science is still failing to provide detailed explanations of cities and their processes. He argues that there is often a naivety amongst those who do not have a background in urban thinking and policy with respect to framing cities and devising solutions, overly focusing on technology and engineering interventions and failing to heed lessons from the long history of urban policy and planning. As he notes, there are no easy solutions to the intractable problems of cities, and urban science will produce no silver bullets, though that is not to say that it will not produce useful insights. He also cautions against the search for universal laws, arguing that urban systems are too large, complex, fluid, and diverse, instead promoting a more tempered approach of understanding individual systems and recognizing, rather than dismissing, the value in other approaches to understanding cities.

It is also the case that scientific analysis is heavily dependent on data quality and contextual information. While urban big data undoubtedly provide numerous opportunities to examine particular systems and issues, they also have a number of limitations. For example, with respect to urban transportation data, while the datasets are rich in volume, they often have limited demographic context – we might know the journeys, but not who took them or why (Batty 2013b). In many cases, the data are being repurposed having been generated by commercial entities for their specific needs but not scientific research. There are thus questions concerning the extent to which repurposed big data provide adequate, rigorous, and reliable surrogates for more targeted, sampled data and how representative such data are of phenomena and populations (Struijs *et al.* 2014). Moreover, big data might seek to be exhaustive, but as with all data they are both a representation and a sample. For example, social media data only relate to those who use a service and are stratified by social class and age, and also include many anonymous and bot accounts. What big data are captured by a system is shaped by: the field of view/sampling frame (where data capture devices are deployed and their settings/parameters); the technology and platform used (different surveys, sensors, lens, textual prompts, and layouts all produce variances and biases in what data are generated); the context in which data are generated (unfolding events mean data are always situated with respect to circumstance); the data ontology employed (how the data are calibrated and classified); and the regulatory environment with respect to privacy, data protection, and security (Kitchin 2014).

Further, much big data have little methodological transparency concerning how they were produced and processed (especially those generated by companies); few metadata with respect to relevance, credibility, timeliness, accessibility, interpretability, coherence, and veracity (accuracy, fidelity, including details of uncertainty, error, bias, reliability, and calibration); and minimal documentation concerning the provenance and lineage of a dataset. And yet it is generally acknowledged that big data can be full of dirty, gamed, and faked data, as well as data being absent (Kitchin 2014). While some might argue that 'more trumps better' and that big data does not need the same standards of data quality, veracity, and lineage because the exhaustive nature of the dataset removes sampling biases and compensates for any errors or gaps or inconsistencies in the data (Mayer-Schonberger and Cukier 2013), it is still the case that garbage-data-in produces garbage-analysis-out.

Urban science and its use of urban big data also raise a number of ethical questions that so far have received little consideration. Since much urban big data are exhaustive and indexical, they raise concerns with respect to privacy, dataveillance and geosurveillance, social sorting, and anticipatory governance (Graham 2005; Kitchin 2016). Many cities are saturated with remote controllable digital CCTV cameras whose footage is increasingly analysed using facial, gait, and automatic number plate recognition software using machine vision algorithms, enabling individuals to be tracked (Kitchin 2016). Smartphones and their apps continuously communicate their location and share with third parties in order to create user profiles and practice targeted location-based advertising (Leszczynski 2018). In a number of cities, for example London and Chicago, sensor networks have been deployed across street infrastructure such as bins and lampposts to capture and track unique phone identifiers such as MAC addresses (Kitchin 2016). Big data analysis can thus reveal highly detailed patterns of spatial behaviour from which other insights, such as mode of travel, activity, lifestyle, co-travellers, can be inferred. The consequence is that individual privacy is eroded with people no longer lost in the crowd and it becomes possible to produce and predict detailed individual and place profiles. These profiles can be used to socially sort and redline populations or to socially sort places with respect to policy interventions. For example, a number of US police forces are now using predictive analytics rooted in urban science research to anticipate the location of future crimes and to direct police officers to increase patrols in those areas and to try and identify potential criminals (Jefferson 2018).

Smart city technologies, the data they generate, and the urban analytics applied to them thus have significant direct and indirect impact on people's everyday lives. Few of those whose data has fed into creating predictive profiles imagined that their data were going to be repurposed to social sort or regulate or control them, or nudge them towards certain behaviours. Generally, these studies – both in universities and industry R&D labs – circumvent notice and consent issues, as well as Institutional Research Boards ethics procedures, by anonymizing and aggregating the data. Nonetheless, the research being undertaken can have effects on those who are unwittingly participating by feeding back into the formulation and delivery of services. In other cases, studies ignore ethical procedures altogether, arguing that data in the public domain (e.g., social media data) are open to carte blanche analysis or that they are entitled to experiment on their own systems without user consent (Kitchin 2016).

To address some of these concerns, some have suggested reconceptualizing cities within urban science and reframing its epistemology (Kitchin 2016). With respect to the first, rather than being cast as bounded, knowable and manageable systems that can be captured, modelled, steered, and controlled in mechanical, linear ways, it is suggested cities need to be understood as fluid, open, complex, multi-level, contingent, and relational systems that are full of culture, politics, competing interests, and wicked problems (Kitchin 2016). With regards to the latter, it is proposed to shift the epistemology towards those employed in critical Geographic Information Science and radical statistics. These approaches employ quantitative techniques, inferential statistics, modelling, simulation, and visual analytics whilst being mindful and open with respect to their shortcomings, drawing on critical social theory to frame how the research is conducted, how sense is made of the findings, and how the knowledge employed (Kitchin 2014; Kitchin *et al.* 2015). Here, it is recognized that there is an inherent politics pervading the datasets analysed, the research conducted, and the interpretations made. Moreover, such a reframing does not foreclose complementing computational social science with 'small data' studies that provide additional and amplifying insights (Crampton *et al.* 2013). In addition, researchers – whether in the public or private domain – need to consider the ethical implications of their work and the uses to which their research is being deployed. Beyond complying with relevant laws and institutional review board (IRB) requirements, urban science practitioners should have a duty of care to citizens not to expose them to harm through its analysis (Kitchin 2016).

Conclusion

Building on earlier rounds of quantitative social science, geocomputation and natural science research, and extending them through the use of new data analytics to extract insights from urban big data, urban science has grown rapidly over the past decade. With the trend in creating smart cities, the on-going growth in the production of urban big data, and large-scale investment in urban science research, this expansion is likely to continue for some time. It is unlikely, however, that urban science will become a new paradigm, producing an integrative approach that replaces the diverse philosophical traditions within urban studies. This is because urban studies continues to produce useful and insightful research and the inherent weaknesses in the epistemology of urban science. Instead, urban science will provide a complementary approach to urban studies and urban informatics and its epistemology is likely to shift and fracture in the same manner as GIScience.

While one could argue that a better approach to city development might be achieved through abandoning urban science and smart cities, others suggest that they are recast in political and epistemological terms. In other words, rather than advocating against the smart city and urban science per se, it is argued that how they are presently conceived and practiced is transformed to be more contextual, relational, and contingent in orientation. Whether such recasting occurs or not, because urban science potentially provides technical, computational solutions to urban problems it will continue to be seen as a viable and profitable means of making sense of cities and creating new products through the analysis of urban big data. In turn, it will continue to influence the development of urban policy and planning and the rollout of smart city initiatives for the foreseeable future.

Acknowledgements

The research for this chapter was funded by an ERC Advanced Investigator award (ERC-2012-AdG 323636-SOFTCITY) and a Science Foundation Ireland grant, Building City Dashboards (15/IA/3090). The chapter in part draws on Kitchin (2016). Thanks to Claudio Coletta, Paolo Cardullo, Liam Heaphy, and Sung-Yueh Perng for comments on the initial draft.

References

Anderson, C. (2008) The end of theory: The data deluge makes the scientific method obsolete. *Wired*, 23rd June. See www.wired.com/2008/06/pb-theory/ (last accessed 2 January 2017).

Batty, M. (2013a) *The New Science of Cities*. Cambridge, MA: MIT Press.

Batty, M. (2013b) *Urban Informatics and Big Data*. London: CASA, University College London. www.spatialcomplexity.info/files/2015/07/Urban-Informatics-and-Big-Data.pdf last accessed 2 Jan 2017.

Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G. and Portugali, Y. (2012) Smart cities of the future. *European Physical Journal Special Topics* 214: 481–518.

Crampton, J.W. (2010) *Mapping: A Critical Introduction to Cartography and GIS*. Malden, MA: Wiley-Blackwell.

Crampton, J.W., Graham, M., Poorthuis, A., Shelton, T., Stephens, M. Wilson, M.W. and Zook, M. (2013) Beyond the Geotag: Situating 'big data' and leveraging the potential of the geoweb. *Cartography and Geographic Information Science* 40(2): 130–139.

Evans, J., Karvonen, A. and Raven, R. (eds) (2016) *The Experimental City*. London: Routledge.

Flood, J. (2011) *The Fires: How a Computer Formula, Big Ideas, and the Best of Intentions Burned Down New York City and Determined the Future of Cities*. New York: Riverhead.

Foth, M. (2009) Preface. In Foth, M. (ed.) *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*, Hershey, PA: Information Science Reference, pp. xxviii–xxxi.

Graham, S.D.N. (2005) Software-sorted geographies. *Progress in Human Geography* 29(5): 562–580.

Jefferson, B.J. (2018) Predictable policing: Predictive crime mapping and geographies of policing and race. *Annals of the American Association of Geographers* 108(1): 1–16.

Kelling, S., Hochachka, W., Fink, D., Riedewald, M., Caruana, R., Ballard, G. and Hooker, G. (2009) Data-intensive science: A new paradigm for biodiversity studies. *BioScience* 59(7): 613–620.

Kitchin, R. (2014) *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.

Kitchin, R. (2015) Positivistic geography. In Aitken, S. and Valentine, G. (eds) *Approaches to Human Geography*, 2nd ed. London: Sage, pp. 23–34.

Kitchin, R. (2016) The ethics of smart cities and urban science. *Philosophical Transactions A* 374(2083): 1–15.

Kitchin, R., Lauriault, T.P. and McArdle, G. (2015) Knowing and governing cities through urban indicators, city benchmarking & real-time dashboards. *Regional Studies, Regional Science* 2: 1–28.

Koonin, S.E. and Holland, M.J. (2013) The value of big data for urban science. In Lanes, J., Stodden, V., Bender, S. and Nissenbaum, H. (eds) *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, Cambridge: Cambridge University Press, pp. 137–152.

Leszczynski, A. (2018) Geoprivacy. In Kitchin, R., Lauriault, T. and Wilson, M. (eds) *Understanding Spatial Media*, London: Sage, pp. 239–248.

Mattern, S. (2013) Methodolatry and the art of measure: The new wave of urban data science. *Design Observer: Places*. 5 November. https://placesjournal.org/article/methodolatry-and-the-art-of-measure/ (last accessed 2 Jan 2017).

Mayer-Schonberger, V. and Cukier, K. (2013) *Big Data: A Revolution that will Change How We Live, Work and Think*. London: John Murray.

Miller, H.J. (2010) The data avalanche is here. Shouldn't we be digging? *Journal of Regional Science* 50(1): 181–201.

O'Sullivan, D. and Manson, S.M. (2015) Do physicists have geography envy? And what can geographers learn from it? *Annals of the Association of American Geographers* 105(4): 704–722.

Pickles, J. (ed.) (1994) *Ground Truth: The Social Implications of Geographic Information Systems*. New York: Guildford Press.

Solecki, W., Seto, K.C. and Marcotullio, P.J. (2013) It's time for an urbanization science. *Environment: Science and Policy for Sustainable Development* 55(1): 12–17.

Struijs, P., Braaksma, B. and Daas, P.J.H. (2014) Official statistics and big data. *Big Data & Society* 1(1): 1–6.

Townsend, A. (2013) *Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia*. New York: W.W. Norton & Co.

Townsend, A. (2015) Making sense of the new science of cities. Rudin Center for Transportation Policy & Management and Data and Society Research Institute, New York University. www.citiesofdata.org/wp-content/uploads/2015/04/Making-Sense-of- the-New-Science-of-Cities-FINAL-2015.7.7.pdf (last accessed 2 Jan 2017).

Wyly, E. (2014) Automated (post)positivism. *Urban Geography* 35(5): 669–690.