4 Data-driven urbanism

Rob Kitchin

Introduction

There is a rich history of data being generated about cities concerning their form, their citizens, the activities that take place, and their connections with other locales. These data have been generated in a plethora of different ways, including audits, cartographic surveying, interviews, questionnaires, observations, photography and remote sensing, and are quantitative and qualitative in nature, stored in ledgers, notebooks, albums, files, databases and other media. Data about cities provide a wealth of facts, figures, snapshots and opinions that can be converted into various forms of derived data, transposed into visualizations, such as graphs, maps and infographics, analysed statistically or discursively, and interpreted and turned into information and knowledge. As such, urban data form a key input for understanding city life, solving urban problems, formulating policy and plans, guiding operational governance, modelling possible futures and tackling a diverse set of other issues. For as long as data have been generated about cities then, various kinds of data-informed urbanism have been occurring.

A new era is, however, presently unfolding wherein data-informed urbanism is increasingly being complemented and replaced by data-driven urbanism. Here, urban operational governance and city services are becoming highly responsive to a form of networked urbanism in which big data systems are prefiguring and setting the urban agenda and are influencing and controlling how city systems respond and perform. In short, we are moving into an era where cities are becoming ever more instrumented and networked, their systems interlinked and integrated, and the vast troves of data being generated used to manage and control urban life. Computation is now routinely being embedded into the fabric and infrastructure of cities producing a deluge of contextual and actionable data which can be processed and acted upon in real-time. Moreover, data that used to be the preserve of a single domain are increasingly being shared across systems enabling a more holistic and integrated view of city services and infrastructures. As such, cities are becoming knowable and controllable in new dynamic ways, responsive to the data generated about them (Kitchin *et al.* 2015). I thus argue that data-driven urbanism is the key mode of production for what have widely been termed smart cities.

In this chapter I provide a critical overview of data-driven urbanism focusing in particular on the relationship between data and the city, rather than network infrastructure, computational or urban issues. The chapter starts by setting out how cities are being instrumented and captured as big urban data, how these data are being used to manage and control cities, and how data-driven urbanism is underpinning the emergence of smart cities. This is then followed by a critical examination of a number of problematic issues related to data-driven urbanism, including: the corporatization of governance (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). More technical data issues such as data quality, the veracity of urban data models and data analytics, and data integration and interoperability are discussed in Chapters 9 and 10.

Big data and smart cities

Since the start of computing era urban data have been increasingly digital in nature, either digitized from analogue sources (manually entered or scanned) or born digital, generated by digital devices, stored as digital files and databases, and processed and analysed using various software systems such as information management systems, spreadsheets and stats packages, and geographic information systems. From the 1980s onwards, public administration records, official statistics and other forms of urban data were released predominately in digital formats and processed and analysed through digital media. However, these data were (and continue to be) generated and published periodically and often several months after generation.

In cases such as exhaustive datasets – for example, detailed framework mapping data or national censuses – new surveys are very infrequent (e.g. 10 years for censuses) and their publication might be 18–24 months after collection, and longer for specific subsets. For domain specific issues, such as transport and traffic flows or public transportation usage, surveys are conducted every few years, using a limited spatial and temporal sampling framework (selected locations for a short period of time). Only a handful of datasets are published monthly (e.g. unemployment rates) or quarterly (e.g. GDP), with most being updated annually due to the effort required to generate them. These data typically have poor spatial resolution, referring to large regions or a nation, and have little disaggregation (e.g. by population classes or economic sectors). In cases where data generation is more frequent, such as remote sensing, only occasional snapshots are bought by city administrations due to their licensing costs. In other cases, such as consumer purchasing (as evidenced in credit card transactions), data have largely been black-boxed within financial institutions. In other words, whilst there has been a range of urban digital data available to urban managers and policy-makers from the 1970s through to 2000s, along with increasingly sophisticated software such as GISs to make sense of them, sources of data were temporally, spatially and domain (scope) limited.

Post-Millennium, the urban data landscape has been transformed, with a massive step-change in the nature and production of urban data, transitioning from small data to big data, wherein the generation of data is continuous, exhaustive to a system, fine-grained, relational and flexible across a range of domains (Kitchin 2014a). From a position of relative data scarcity, the situation is turning to one of data deluge. This is particularly the case with urban operational data wherein traditional city infrastructure, such as transportation (e.g. roads, rail lines, bus routes, plus the vehicles/carriages) and utilities (e.g. energy, water, lighting), have become digitally networked, with grids of embedded sensors, actuators, scanners, transponders, cameras, meters and GPS (constituting what has been called the Internet of Things) producing a continuous flow of data about infrastructure conditions and usage. Many of these systems are generating data at the individual level, tracking travel passes, vehicle number plates, mobile phone identifiers, faces and gaits, buses/trains/taxis, meter readings, etc. (Dodge and Kitchin 2005). These are being complemented with big data generated by: (a) commercial companies such as mobile phone operators (location/movement, app use, activity), travel and accommodation sites (reviews, location/movement, consumption), social media sites (opinions, photos, personal info, location/movement), transport providers (routes, traffic flow), website owners (clickstreams), financial institutions and retail chains (consumption, in-store movement, location), and private surveillance and security firms (location, behaviour) that are increasingly selling and leasing their data through data brokers, or making their data available through APIs (e.g. Twitter and Foursquare); (b) crowdsourcing (e.g. Open Street Map) and citizen science (e.g. personal weather stations) initiatives, wherein people collaborate on producing a shared data resource or volunteer data. Other kinds of more irregular urban big data include digital aerial photography via planes or drones, or spatial video, LiDAR (light detection and ranging), thermal or other kinds of electromagnetic scans of environments that enable the mobile and real-time 2D and 3D mapping of landscapes. And whilst official statistics are largely still waiting to undergo the data revolution (Kitchin 2015), the generation of public administration data has been transformed through the use of e-government online transactions that produce digital data at the point-of-collection.

We are at start of this new big data era and the flow and variety of urban data is only going to grow and diversify. Moreover, whilst much of these data presently remain in silos and are difficult to integrate and interlink due to varying standards and formats, they will increasingly be corralled into centralized systems such as inter-agency control rooms for monitoring the city as a whole or what have been termed city operating systems. With regards to the former, the Centro De Operacoes Prefeitura Do Rio in Rio de Janeiro, Brazil, a data-driven city operations centre pulls together into a single location real-time data streams from 30 agencies, including traffic and public transport, municipal and utility services, emergency and security services, weather feeds, information generated by employees and the public via social media, as well as administrative and statistical data, and is overseen by a staff of 180 data operatives (see Figure 4.1 for two examples of urban control rooms). City operating systems are effectively Enterprise Resource Planning (ERP) systems designed to coordinate and operate the activities of large companies repurposed for cities. Examples include Microsoft's *CityNext*, IBM's *Smarter City*, Urbiotica's *City Operating System* and PlanIT's *Urban Operating System*. With the advent of the open data movement some of these data also feed into public-facing urban dashboards that provide a mix of interactive visualizations of real-time, public administration and official statistical data (Kitchin *et al.* 2015; see Chapter 10).

Further, the production of these new big data has been accompanied by a suite of new data analytics designed to extract insight from very large, dynamic datasets, consisting of four broad classes: data mining and pattern recognition; data visualization and visual analytics; statistical analysis; and prediction, simulation and optimization (Miller 2010; Kitchin 2014b). These analytics rely on machine learning (artificial intelligence) techniques and vastly increased computational power to process and analyse data. Moreover, they enable a new form of data-driven science to be deployed that rather than being theory-led seeks to generate hypotheses and insights 'born from the data' (Kelling *et al.* 2009). This is leading to the development of 'urban informatics' (Foth 2009), an informational and human–computer interaction approach to examining and communicating urban processes, and 'urban science', a computational modelling approach to understanding and explaining city processes that builds upon and radically extends quantitative forms of urban studies that have been practised since the 1950s, blending in geocomputation, data science and social physics (Batty 2013). Whereas urban informatics is more human-centred, interested in understanding and facilitating the interactions between people, space and technology, urban science promises to not only make sense of cities as they presently are (by identifying relationships and urban 'laws'), but to also predict and simulate likely future scenarios under different conditions, potentially providing city managers with value insight for planning and development decision-making and policy formulation.

Figure 4.1 Urban control rooms: (a) Rio de Janeiro, (b) Dublin

Urban big data, city operating systems, urban informatics and urban science analytics provide the basis for a new logic of urban control and governance – data-driven urbanism – that enables real-time monitoring and steering of urban systems and the creation of what has widely been termed smart cities. The notion of a smart city can be traced back to experiments with urban cybernetics in the 1970s (Flood 2011; Townsend 2013), the development of new forms of city managerialism and urban entrepreneurship, including smart growth and new urbanism, in the 1980s and 1990s (Hollands 2008; Wolfram 2012; Söderström *et al.* 2014; Vanolo 2014), and the fusing of ICT and urban infrastructure and development of initial forms of networked urbanism from the late 1980s onwards (Graham and Marvin 2001; Kitchin and Dodge 2011). As presently understood, a smart city is one that strategically uses networked infrastructure and associated big data and data analytics to produce a:

- *smart economy* by fostering entrepreneurship, innovation, productivity, competitiveness, and producing new forms of economic development such as the app economy, sharing economy and open data economy;
- *smart government* by enabling new forms of e-government, new modes of operational governance, improved models and simulations to guide future development, evidence-informed decision-making, better service delivery, and making government more transparent, participatory and accountable;
- *smart mobility* by creating intelligent transport systems, efficient inter-operable multi-modal public transport, smart parking and sharing services related to taxis and bikes;
- *smart environments* by promoting sustainability and resilience and the development of green energy;
- *smart living* by improving quality of life, increasing safety and security and reducing risk;
- *smart people* by creating a more informed citizenry and fostering creativity, inclusivity, empowerment and participation (Giffinger *et al.* 2007; Cohen 2012).

In short, the smart city promises to solve a fundamental conundrum of cities – how to reduce costs and create economic growth and resilience at the same time as producing sustainability and improving services, participation and quality of life – and to do so in common-sense, pragmatic, neutral and supposedly apolitical ways by utilizing a fast-flowing torrent of urban data and data analytics, algorithmic governance and responsive networked urban infrastructure. Moreover, much more information is being placed into the hands of the public to aid decision-making, navigation and participation through a plethora of locative social media (apps that tell them about the city and which they can contribute to), open data sites, public dashboards, hackathons and so on.

The notion of smart cities, and the mode of data-driven urbanism, have not however been universally welcomed and have been subject to a number of critiques. First, smart city initiatives treat cities as a set of knowable and manageable systems that act in largely rational, mechanical, linear and hierarchical ways and can be steered and controlled, rather than dealing with cities as complex, messy, contingent cities full of wicked problems (Kitchin *et al.* 2015). Second, smart city initiatives are largely ahistorical, aspatial and homogenizing in their orientation and intent, treating cities as if they are all alike in terms of their political economy, culture and governance (Greenfield 2013). Third, an emphasis is placed on creating technical rather political/social/policy solutions to urban problems thus overly promoting technocratic forms of governance (Morozov 2013). Fourth, the project of producing smart cities tends to reinforce existing power geometries and social and spatial inequalities rather than eroding or reconfiguring them (Datta 2015). Fifth, the approach fails to recognize the politics of urban data and the ways in which they are the product of complex socio-technical assemblages (Kitchin 2014b). Sixth, the smart city agenda is being overly driven by corporate interests who are using it to capture government functions as new market opportunities rather than serve a public good (Hollands 2008). Seventh, networking city infrastructure potentially creates buggy, brittle and hackable urban systems (Kitchin and Dodge 2011; Townsend 2013). And finally, data-driven urbanism produces a number of activities that have profound social, political, ethical consequences, including dataveillance and extensive geosurveillance, social and spatial sorting, and anticipatory governance (Graham 2005; Kitchin 2014a).

In the rest of this chapter, I want to concentrate on the last four critiques, and in particular their associated data issues (rather than other aspects of the technical stacks of urban socio-technical assemblages, and wider political-economic framing and effects) as way of further illustrating some of the challenges posed by data-driven urbanism and the need to further examine the relationship between data and the city.

Data and the city

The politics of urban data

One of the key arguments for adopting a data-driven approach to urban governance is that it provides a strong evidence-based approach to decision-making, system control and policy formation, rather than one that is anecdotal, clientelist or localist. A data-driven approach, it is argued, is less susceptible to political influence and instead is driven by objective, neutral facts in a technocratic, common-sense, pragmatic way. Technical systems and the data they produce are objective and non-ideological and thus politically benign. Sensors, networked infrastructure and computers it is contended have no inherent politics – they simply measure a value, communicate those values, and process, analyse and display the data using scientific principles, thus producing measurements, records and information that reflect the truth about cities. And while data from social systems, such as social media platforms (e.g. Twitter), are inherently more subjective and noisy, they provide a direct reflection of the views, interactions and behaviour of people, in contrast to official surveys which reflect what people say they do or think (or what they think the surveyor wants to hear), thus providing better ground truthing of social reality. As such, big data about cities can be taken at face value and used unconditionally shed light on cities and to manage and control urban systems and infrastructure and guide urban policy.

The reality is somewhat different for two reasons. First, there are a number of technical issues concerning data coverage, access and quality that means that the view data presents of the city is always partial and subject to caution (see Chapters 2, 3, 10, 15, 17). Second, data are the products of complex socio-technical assemblages that are framed and shaped by a range of technical, social, economic and political forces and are designed to produce particular outcomes (Kitchin 2014b; see Figure 4.2). On the one hand, what data are produced, how they are handled, processed, stored, analysed and presented is the result of a particular technical configuration and how it is deployed (e.g. where sensors are located, their field of view, their sampling rate, their settings and calibration, etc.). On the other hand, how a system is designed and run is influenced by systems of thought, technical know-how, the regulatory environment, funding and resourcing, organizational priorities and internal politics, institutional collaborations and marketplace demand. In other words, a data assemblage possesses a '*dispositif*', defined by Foucault (1980 [1977]: 194) as a: 'thoroughly heterogeneous ensemble consisting of discourses, institutions, architectural forms, regulatory decisions, laws, administrative measures, scientific statements,

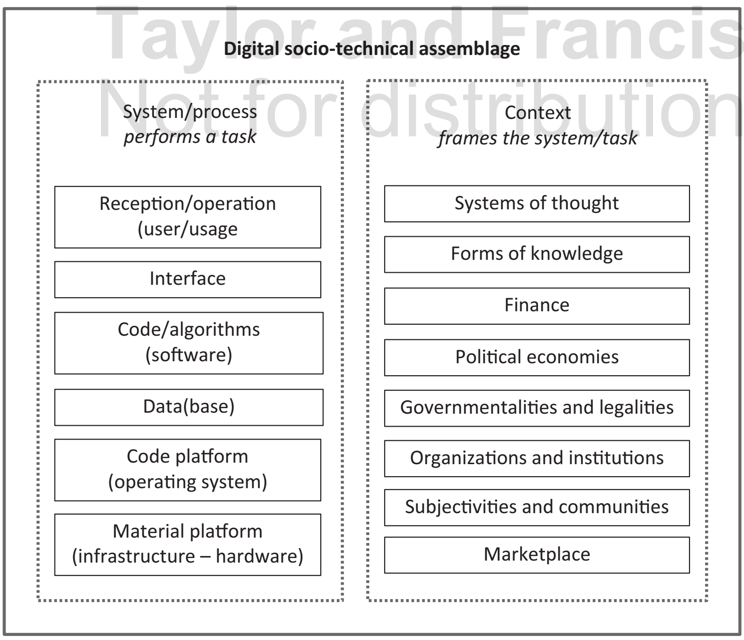

Figure 4.2 A data assemblage

philosophical, moral and philanthropic propositions'. For Foucault, a *dispositif* is inherently political producing what he terms 'power/knowledge', that is knowledge that fulfils a strategic function. In other words, urban big data are never neutral and objective, but rather are situated, contingent, relational and framed and used contextually to try and achieve certain aims and goals (to monitor, enhance, empower, discipline, regulate, control, produce profit, etc.). Or to put it another way, urban data are never raw but are always already cooked to a particular recipe for a particular purpose (Bowker 2005; Gitelman 2013). As such, data-driven urbanism is thoroughly political, seeking to produce a certain kind of city. It is thus necessary when examining urban big data to critically unpack their associated data assemblage (including the entire technical stack – infrastructure, platform, software/algorithms, data, interface) to document how it is constituted and works in practice to produce urban processes and formations, and for whose benefit.

Data access, data ownership and data control

As already noted, much of the data presently being generated about cities are produced by commercial companies, such as mobile phone operators, and private utility and transport companies. For them, their data are a valuable commodity that provides competitive advantage or an additional revenue stream if sold/leased, and they are under no obligation to share freely the data they generate through their operations with city managers or the public. As noted in 2014 by the British Minister for Smart Cities, Dan Byles MP,[1] the privatization of public services in the UK and elsewhere has also meant the privatization of their associated data unless special provision was made to ensure it was shared with the city or made open. Similarly, access to data within public–private partnerships and semi-state agencies, or state agencies operating as trading funds (such as the Met Office and Ordnance Survey in the UK who generate significant operating costs by selling data and services), can be restricted or costly to purchase. Consequently, key framework datasets (e.g. detailed maps) can have limited access and data concerning transportation (e.g. bus, rail, bike share schemes, private tolls), energy and water be entirely black-boxed. Even within the public sector, data can be siloed within particular departments and not be shared with other units within the organization, or be open for other institutions or the public to use. As such, whilst there might be a data revolution underway, access to much of that data is limited, and there are a number of issues that need to be explored with respect to data ownership and data control, especially with respect to procurement and the outsourcing or privatization of city services. Moreover, even if all data were to be open and shared it needs to be acknowledged that there are still many aspects about cities where data generation is weak or absent. For example, in a recent audit of Dublin datasets to determine whether the city was in a position to apply for ISO37120 (the ISO standard for city indicators) data could only be sourced for 11 of 100 indicators sought (predominately because the data sought was either privatized or released at an inappropriate scale).

Data security and data integrity

One of the prime anxieties of networking infrastructure and ubiquitous urban computing is the creation of systems and environments which are inherently buggy and brittle and are prone to viruses, glitches, crashes and security hacks (Kitchin and Dodge 2011; Townsend 2013). As Mims (2013) notes, any networked device is open to be hacked and its data stolen and used for criminal purposes, or corrupted, or controlled remotely, or misdirected, or to spy on its users. The media report almost daily on large-scale data breaches of commercial companies and state agencies and the theft of valuable personal data, with several incidents of city infrastructure such as traffic management systems being hacked, disabled and controlled (Paganini 2013). As Townsend (2013) notes, the notion of smart cities takes two open, highly complex and contingent systems – cities and computing – and binds and networks them together, meaning that data-driven, networked urbanism has in-built vulnerabilities. Moreover, as urban systems evolve to become more complex, interconnected and interdependent these vulnerabilities potentially multiply (Townsend 2013; Kitchin 2016). Creating secure big urban data systems is thus set to be a significant ongoing task if public trust in their purported benefits are to be gained and maintained. Another significant element in upholding trust is how and to what purposes the data are deployed.

Data uses and ethics

Urban big data are presently being used to undertake a diverse range of tasks, some of which seem relatively benign, such as monitoring city lighting with the aim of improving the quality of light and reducing its cost, and others more clearly political, such as directing policing activity. A significant concern is that as more and more data about cities and their citizens are generated, privacy becomes eroded (Kitchin 2016). Privacy is considered a basic human right, a condition that people expect and value in developed countries. Yet, as sensors, cameras, smartphones and other embedded and mobile devices generate ever more data it becomes increasingly difficult to protect, with individuals leaving ever greater quantities of digital footprints (data they themselves leave behind) and data shadows (information about them generated by others). Such troves of data are amenable to dataveillance, a mode of surveillance enacted through sorting and sifting datasets in order to identify, monitor, track, regulate, predict and prescribe (Clarke 1988; Raley 2013), and geosurveillance, the tracking of location and movement of people, vehicles, goods and services and the monitoring of interactions across space (Crampton 2003; see Table 4.1). Given the always-on nature of many of these systems, and the tracking of unique identifiers, such dataveillance and geosurveillance are becoming continuous and fine-grained with, for example, mobile phone companies always knowing the location of a phone (Dodge and Kitchin 2005). Moreover, as data minimization norms become relaxed there are anxieties that data are being shared, combined and used for purposes for which they were never intended (Kitchin 2014b). In particular, the last 20 years have witnessed the rapid growth of a number of data brokers who capture, gather together and repackage data for rent (for one time use or use under licensing conditions) or re-sale, and produce various derived data and data analytics (CIPPIC 2006).

Whilst focusing on different markets, data brokers seek to mesh together offline, online and mobile data to provide comprehensive views of people and places and to construct personal and geodemographic profiles (Goss 1995; Harris *et al.* 2005). These profiles are then used to predict behaviour and the likely value or worth of an individual and to socially sort them with respect to credit, employment, tenancy and so on (Graham 2005). The concern is that these firms practice a form of 'data determinism' in which individuals are not profiled and judged just on the basis of what they have done, but on the prediction of what they might do in the future using algorithms that are far from perfect, and yet are black-boxed and lack meaningful oversight and remediate procedures (Ramirez 2013). Such anticipatory governance can have far reaching effects. For example, a number of US police forces are now using predictive analytics to anticipate the location of future crimes and direct patrols, and to identify individuals most likely to commit a crime in the future, designating them pre-criminals (Stroud 2014). In such cases, a person's digital footprints and data shadow does more than follow them; it precedes them. Data assemblages then do not act as cameras reflecting the world as it is, but rather as engines shaping the world in diverse ways (Mackenzie 2008).

Table 4.1 Movement and location tracking. Compiled from Kitchin (2016)

Remote controllable digital CCTV cameras	Can zoom, move and track individual pedestrians and vehicles. Analysis and interpretation increasingly aided by facial, gait and automatic number plate recognition (ANPR) using machine vision algorithms.
Smartphone phone tracking	Location is communicated to telecommunications providers through the cell masts connected to the sending of GPS coordinates, or connections to wifi hotspots.
Sensor networks	Sensors deployed on street infrastructure such as bins and lampposts or in shops/malls capture and track phone identifiers such as MAC addresses.
Wifi meshes	The IDs of devices which access or try to access a wifi network are captured and tracked between wifi points.
Smart card tracking	Barcodes, magnetic strips or embedded RFID chips are tracked when they are scanned to gain entry to buildings or transportation.
Active GPS tracking	Embedded GPS in devices and vehicles communicate location and movement via cellular or satellite networks.
Transponder tracking	Transponders with embedded RFID chips broadcast their IDs and are tracked by scanning receivers, commonly used in automatic road tolling or electronic tagging of people on probation.
Other staging points	Such as using ATMs, credit cards or checking a book out of a library that leaves a digital record.

Conclusion

We are entering an era where computation is being routinely embedded into urban environments and networked together, and people are moving about with smartphones that ensure always available connectivity and access to information. These devices and infrastructures are producing and distributing vast quantities of data in real-time, and they are also responsive to these data and the analytics undertaken on them enabling new kinds of monitoring, regulation and control. Cities then are becoming data-driven and are enacting new forms of algorithmic governance. However, the data and algorithms underpinning them are far from objective and neutral, but rather are political, imperfect and partial. The smart cities that data-driven, networked urbanism purports to create are then smart in a qualified sense. Their production and operation is based on much more data and derived information than previous generations of urbanism, but it is a form of urbanism that is nonetheless still selective, crafted, flawed, normative and politically inflected. Moreover, while the instrumental rationality of data-driven, networked urbanism promotes urban knowledge and management rooted in a quite narrowly framed 'episteme (scientific knowledge) and teche (practical instrumental knowledge)', it is important that other forms of knowing, such as 'phronesis (knowledge derived from practice and deliberation) and metis (knowledge based on experience)' (Parsons 2004: 49) are not silenced, providing both a counter-weight to the limits of smart cities and positions from which to reflect on, critique and recast the production of data-driven urbanism. Indeed, whilst data-driven urbanism undoubtedly provides a set of solutions for urban problems, we also have to recognize that it has a number of shortcomings and a number of potential perils. The challenge facing urban managers and citizens in the age of smart cities is to realise the benefits of planning and delivering city services using a surfeit of data, evidence and real-time responsive systems whilst minimizing any pernicious effects. To do that we have to be as smart about data and data analytics as we would like to be about cities.

Acknowledgements

The research for this chapter was provided by a European Research Council Advanced Investigator Award, 'The Programmable City' (ERC-2012-AdG-323636).

Note

1 www.youtube.com/watch?v=3E3RpGMKbhg.

References

Batty, M. (2013) *The New Science of Cities*. Cambridge, MA: MIT Press.
Bowker, G. (2005) *Memory Practices in the Sciences*. Cambridge, MA: MIT Press.
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. Available from: https://cippic.ca/sites/default/files/May1-06/DatabrokerReport.pdf [accessed 6 February 2017].

Clarke, R. (1988) 'Information technology and dataveillance', *Communications of the ACM* 31(5): 498–512.

Cohen, B. (2012) 'What exactly is a smart city?', *Fast Co.Exist*, 19 September, available from: www.fastcoexist.com/1680538/what-exactly-is-a-smart-city [accessed 6 February 2017].

Crampton, J. (2003) 'Cartographic rationality and the politics of geosurveillance and security', *Cartography and Geographic Information Science* 30(2): 135–148.

Datta, A. (2015) 'New urban utopias of postcolonial India: 'Entrepreneurial urbanization' in Dholera smart city, Gujarat', *Dialogues in Human Geography* 5(1): 3–22.

Dodge, M. and Kitchin, R. (2005) 'Codes of life: Identification codes and the machine-readable world', *Environment and Planning D: Society and Space* 23(6): 851–881.

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. (ed.) (2009) *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*. New York: IGI Global.

Foucault, M. (1980 [1977]) 'The confession of the flesh', in C. Gordon (ed.), *Power/Knowledge*. New York: Pantheon Books, pp. 194–228.

Giffinger, R., Fertner, C., Kramar, H., Kalasek, R., Pichler-Milanović, N. and Meijers, E. (2007) *Smart cities: Ranking of European medium-sized cities*. Centre of Regional Science, Vienna UT, available from: www.smart-cities.eu/download/smart_cities_final_report.pdf [accessed 12 October 2015].

Gitelman, L. (ed.) (2013) *'Raw Data' is an Oxymoron*. Cambridge, MA: MIT Press.

Goss, J. (1995) '"We know who you are and we know where you live": the instrumental rationality of geodemographics systems', *Economic Geography* 71: 171–198.

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

Graham, S. and Marvin, S. (2001) *Splintering Urbanism: Networked Infrastructures, Technological Mobilities and the Urban Condition*. New York: Routledge.

Greenfield, A. (2013) *Against the Smart City*. New York: Do Publications.

Harris, R., Sleight, P. and Webber, R. (2005) *Geodemographics, GIS and Neighbourhood Targeting*. Chichester: Wiley.

Hollands, R.G. (2008) 'Will the real smart city please stand up?', *City* 12(3): 303–320.

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. (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. (2015) 'The opportunities, challenges and risks of big data for official statistics', *Statistical Journal of the International Association of Official Statistics* 31(3): 471–481.

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 [accessed December 2016].

Kitchin, R. and Dodge, M. (2011) *Code/Space: Software and Everyday Life*. Cambridge, MA: MIT Press.

Kitchin, R., Lauriault, T. and McArdle, G. (2015) 'Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards', *Regional Studies, Regional Science* 2: 1–28.

Mackenzie, D. (2008) *An Engine, Not a Camera. How Financial Models Shape Markets*. Cambridge, MA: MIT Press.

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

Mims, C. (2013) 'Coming soon: the cybercrime of things', *The Atlantic*, 6 August, available from: www.theatlantic.com/technology/archive/2013/08/coming-soon-the-cybercrime-of-things/278409/ [accessed 7 August 2013].

Morozov, E. (2013) *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems That Don't Exist*. New York: Allen Lane.

Paganini, P. (2013) 'Israeli road control system hacked, causes traffic jam on Haifa highway', *The Hacker News*, 28 October, available from: http://thehackernews.com/2013/10/israeli-roadcontrol-system-hacked.html [accessed 13 November 2013].

Parsons, W. (2004) 'Not just steering but weaving: Relevant knowledge and the craft of building policy capacity and coherence', *Australian Journal of Public Administration* 63(1): 43–57.

Raley, R. (2013) 'Dataveillance and countervailance', in L. Gitelman (ed.), *'Raw Data' is an Oxymoron*. Cambridge, MA: MIT Press, pp. 121–146.

Ramirez, E. (2013) 'The privacy challenges of big data: A view from the lifeguard's chair', *Technology Policy Institute Aspen Forum*, 19 August, available from: http//ftc.gov/speeches/ramirez/130819bigdataaspen.pdf [accessed 11 October 2013].

Söderström, O., Paasche, T. and Klauser, F. (2014) 'Smart cities as corporate storytelling', *City* 18(3): 307–320.

Stroud, M. (2014) 'The minority report: Chicago's new police computer predicts crimes, but is it racist?', *The Verge*, 19 February, available from: www.theverge.com/2014/2/19/5419854/the-minority-report-this-computer-predicts-crime-but-is-it-racist [accessed 4 December 2016].

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

Vanolo, A. (2014) 'Smartmentality: The smart city as disciplinary strategy', *Urban Studies* 51(5): 883–898.

Wolfram, M. (2012) 'Deconstructing smart cities: An intertextual reading of concepts and practices for integrated urban and ICT development', in M. Schrenk, V.V. Popovich, P. Zeile and P. Elisei (eds), *Re-Mixing the City: Towards Sustainability and Resilience?* REAL CORP, pp. 171–181.