Batty, M. (2017) Producing Smart Cities, In Kitchin, R., Lauriault, T., and Wilson, M. **Understanding Spatial Media**, Sage Publications, London, in press.

19

PRODUCING SMART CITIES

MICHAEL BATTY

CHANGING CONCEPTIONS OF THE CITY

Just as our cities have become less polluted, greener and, at least in the west, more prosperous, our interest in them has dramatically accelerated. Big cities, in particular, are the flavour of the month. They attract the young and the talented, the cool and the creative. They are fast becoming the places where the future is being invented through new information technologies (IT), new lifestyles, new modes of entertainment, consumption and work. Into this milieu has come the idea that cities can be automated in the same way as industrial systems were automated in earlier generations, but this time around, it is the collective functions of cities that are being recreated through new IT and new ways of communication. Many of these technologies constitute new spatial media given their production, processing and dependence on vast quantities of spatial big data and heavy utilisation of spatial interfaces such as maps and spatialisations. In short, we now speak of the 'smart city', wherein cities are endowed with some intelligence in their functions, which in the past have been somewhat dumb in their operations. In one sense, our new-found interest in cities cannot be separated from developments in digital technologies. Advances in computation and communications are leading to many new forms of social media that in turn are augmenting and changing the functions that continue to dominate the way our cities work. In this sense, the development of the smart city is just another wave in the development of information technologies that have spread out from science to business and government over the past 50 years, and in the past 20 years from the individual to the collectivity, culminating in what some are calling a new 'operating system' for the city.

There is a good deal of hype involved in this notion that our cities are suddenly becoming smart simply because our IT have continued to spread out and diffuse into all corners of our lives. But once we cut through the rhetoric, the smart city essentially embraces the use of digital technologies that are being used to automate routine day-to-day functions that enable populations to communicate with each other and with the artefacts of the built environment that are used to make the city work. Moving people, goods, information and ideas using technologies that complement traditional modes of communication are key, and providing routine government and non-market services using IT such as those related to location, and engaging citizens in various public domains, are central to the smart city. To do this, the entire panoply of IT devices underpins these patterns of delivery and communication. In the past decade, personal devices that we use to communicate with one another and with artefacts have become central to such operations, the exhaust from which is providing enormous quantities of data – big data (vast quantities of which are georeferenced) – that we are beginning to use to make sense of the way these augmented functions are changing the city.

Ways of using these new devices to capture and spread new media are enriching and changing the way we communicate socially, and such social media are providing diverse ways to supplement many traditional functions involving work and entertainment. We are beginning to provide new interfaces to the way cities operate by synthesising these data visually into dashboards that provide an instant picture of city life (see Chapter 7), and for these we are evolving many new forms of urban analytics (see Chapter 14) enabling us to produce more insightful analyses and predictions of how cities can be improved in terms of their functioning. At least that is the intention and the hope. So to summarise, new forms of digital communications, new services delivered digitally and new communities fashioned from social media constitute the smart city, made possible by all-pervasive digital devices that are either embedded into the fabric of our cities producing real-time streams of sensed data or operated in mobile fashion by ourselves. These are all producing data that a new array of analytics is beginning to make sense of, directed towards a new type of science that is in enabling us to better understand cities in general and the smart city in particular.

PROGRESS TOWARDS A NEW SCIENCE

Until quite recently our thinking about cities in formal ways has been based on their development to an assumed equilibrium over quite large time spans. In fact, we tend to articulate cities as we see them at a single point in time. Consequently, many of our theories about how their functions generate the physical forms we see and how their populations behave with respect to the way land is used describe a timeless world in which locations and forms are explained as if they were suddenly created at that instant. Most of our formal models suggest a world in equilibrium in contrast to the very large body of urban studies focusing on explaining the city more descriptively in verbal terms, which does examine its dynamics but again, usually over the longer term. Only very recently has the notion that we can explain the routine short-term dynamics of the city more formally come onto the agenda. This is as much through the emergence of big data as it is through the shortening of our attention spans in a world that is ever faster and seemingly more responsive.

At present, our urban theories provide a mixture of those that deal with the short term and the long term, with cities in equilibrium as well as out-of-equilibrium

(Batty, 2016). There is, however, no sense in which there is an integrated scientific perspective making sense of all this variety. There are, though, some features of what we know about cities that are beginning to forge such a new science. Central to this is the notion that cities are all about flows – about interactions and relationships – which underpin the patterns that we observe with respect to locations, places and the way we carve the city up into different spaces. To understand location, interactions, movements, communications and so on is now deemed to be essential (Batty, 2013). This is a long-standing idea but it is surprising that we have not taken it more to heart. With the rise of social media and with movement and communications being recorded digitally through everything from email to smart cards, the idea of the city as a communications medium has become critical in ways that commentators such as Richard Meier (1962) and Melvin Webber (1964) anticipated so acutely and so long ago.

Data that are streamed or captured by mobile or fixed sensors are largely unstructured, but in cities they tend to show great heterogeneity and diversity. They are usually highly disaggregate if collected and associated with individuals using smart cards or phones. In this sense, they provide rather good material for micro and agent-based simulations. Because they are temporally dynamic in their basic form and usually extensive, they require different kinds of analytics and modelling from that associated with much more structured, smaller datasets that are collected at a cross-section in time from population censuses and the like. Multivariate statistics, which had developed rapidly from over half a century ago, are increasingly inappropriate on which to base such analytics and new forms of data mining focusing on extracting patterns from machine learning are being evolved. Models have become much more disaggregate and new computer methods based on massively parallel distributed processing are now required to make sense of such data. Visualisation is increasingly important and this focus has changed from long-term and cross-sectional simulation to much more routine operations and functions in cities. Disruption, crime and health emergencies, for example, occurring daily are of more concern with respect to modelling and prediction. All of this is being painfully grafted onto the overall structure of knowledge in this domain, while some of the old certainties with respect to how cities are structured are rapidly disappearing. These certainties have gone as a new paradigm focused on complexity has propelled forward formal frameworks associated with agent-based modelling, micro-simulation, serious games, network processes and cellular automata.

It is going to take some time for this new science to be constructed and like all knowledge it is contingent on the times in which it is being invented and discovered. Cities are getting ever more complex and IT which are being fast embedded into their fabric are changing the very way in which we behave. There is now a clear imperative that suggests that as we observe and probe the city, we change it, just as we do in the more esoteric branches of physics dealing with quanta. Indeed, the models and analytics that we are using are changing the very systems that we are seeking to understand and manipulate. In the smart city, everything is contingent and changeable. To give some sense of this, I set out three key dimensions of the future city: integrated platforms or operating systems for cities; communications and movement; crowdsourcing and citizen science (see Chapter 12). All of these are being informed by new IT that are central to how we behave and interact with one another in the smart city.

ADAPTING NEW TECHNOLOGIES

INTEGRATED PLATFORMS AND CITY OPERATING SYSTEMS

The most optimistic perspective on the smart city suggests that we can apply quite literally the idea of the operating system to organising many functions that a city generates. This, to an extent, is fanciful for it assumes we know and understand the range of functions and how they relate to each other and can integrate them somewhat seamlessly into a form where they are interoperable. Our experience hitherto with large IT systems is not good, particularly where strong top-down imperatives are used to build such systems, as for example in national agencies – in Britain the National Health system is typical and indeed apocryphal. Many of the largest IT companies such as IBM, which concentrate on computer services and software, and Cisco, which dominates networks and switching, have implicitly endorsed this idea and in the flurry of planned technological new towns such as Masdar and Songdu, there are demonstrations of how such systems might be developed. In fact, Living PlanIT are developing the 'PlanIT Urban Operating System' originally for an area east of Porto in Portugal, but this is, I feel, more a metaphor for integrating IT than a realisable proposal.

The difficulties of coupling systems do not pertain just to interoperability but to questions of data integration and even the styles of models and simulations that pertain to different sectors. For example, unless there are common keys between datasets – of which there are very few in cities, geocodes being the most obvious – it is impossible to integrate anything. Transport for London can measure demand in real time from its Oyster smart card, which records all tap-ins and tap-outs of its travellers, but it impossible to link this to the supply of trains, which is also known at precise locations and time instants from their operating systems. Within a tube station travellers cannot be tracked – it is illegal – and in any case the Wi-Fi is not sufficiently detailed and may never be good enough to track 'what people' get onto 'what trains'. Currently it is impossible even with state-of-the-art equipment and data to cement this linkage. If Transport for London is unable to do this, no one else can and this kind of circumstance is writ large across our cities.

A more modest aim for cities is integration that is built from the bottom up – in what are increasingly called platforms (Grech, 2015). These are digital forums for interfacing a variety of application programming interfaces (APIs) – pieces of software that can be queried, linked and coupled to each other, but ordered in certain ways that reflect the purpose of the integration. Generic software is being invented in this way so it can take account of very different pieces of established and not-so-established software and data that can be linked into platforms that offer the generic software as a service. One of the big differences between seeing cities as platforms or as operating systems is not just simply one of ambition, but one of modularity where the components of the platform can be incessantly replaced with better and/or different versions of the same, producing the possibility of continual disruption as new software is created from the old that is being destroyed. This model is one that appears to be endemic to IT as witnessed in the great transition from hardware to software, from large to tiny computers and from remote individual access to all pervasive networking. IT are increasingly subject to Schumpeter's notion of 'creative destruction'; no longer within the cyclical Kondratieff wave of 75 years' duration but now being almost continuous in its transformation (Batty, 2015).

To date, no operating systems or platforms have actually been built that embody the notions of the smart city, unless one counts ad hoc attempts at stitching available software and sensor systems together in a loosely coupled manner. The rhetoric is more one of metaphor than of actuality, but what have been developed are portals where data are drawn together, either in dashboard-like interfaces, one of which the Centre for Advanced Spatial Analysis (CASA) has constructed for London (see Figure 19.1) (O'Brien et al., 2014) or in physical locations built around the idea of city control rooms. Most big cities have such rooms for specialist purposes where manual and automatic controls of systems such as water, electricity, traffic and so on are operated. Proposals for city-wide services such as IBM's Rio de Janeiro implementation show what is possible (Kitchin et al., 2015). The basic problem remains, however, that first, the data and control that are possible are not easy to integrate across different systems, notwithstanding the fact that such portals are still of interest in providing a big picture. And second, many of the key indicators that we need to visualise with respect to how cities are performing cannot be sensed automatically for they depend on social and economic data such as housing, ownership and

FIGURE 19.1 The London Dashboard – real-time data pertaining to the city

tenancy, market prices, unemployment and so on. These data are slowly coming but they may never be on the same frequency as more physical and travel-based data for they are only likely to be released on monthly perhaps, at best, weekly cycles. This changes the picture of how a city is performing quite radically from the minute-by-minute displays that characterise the current generation of dashboards.

COMMUNICATIONS AND MOVEMENT IN THE CITY

If you examine the dashboard in Figure 19.1, you will see many inputs of data that pertain to how people, materials and ideas are moving within the city – frequencies of subway (tube) services, what is trending on Twitter and so on. Cities are places where people come together to engage socially and economically, to share and divide their labour, to add value to what they do to sustain themselves. Cities are quintessentially market places for ideas and material transactions. If we could all interact at the same physical location, we would not require any physical movements to enable these transactions to take place, but given the exigencies of physical structure and the fact that as human beings our own personal space is geometrically distinct, many different kinds of movements take place in cities to enable cities to function effectively. Much urban planning attempts to reduce the scale and volume of these movements although this is quite a controversial view in that some argue that movement in fact is necessary for the good functioning of cities and that in the modern world, globalisation essentially projects this idea to its extreme. The argument is one that relates to density, compactness, sprawl and city size, over which there is little agreement.

We have good data on how public transport is being automated in London and this is increasingly the case in many big cities, where smart cards are replacing more manually based charging systems. Oyster Cards are used in London to record where and when and who taps in and out of the subway and other train network stations. Buses are trickier as one-man drivers mean that only tap-in can be monitored. In fact in Singapore, the EZ-Link system records tap-in and tap-out for trains and buses, but this difference reflects cultural mores that are important in understanding the extent to which cities can be automated. This kind of dataset provides just one perspective on physical travel. Because it is unstructured and relates to places where people enter and exit a transit system, it is very much an incomplete picture of how people move from fixed locations that relate to places where they undertake activities such as education, entertainment, work and so on. It needs to be integrated with other datasets so it can be given structure.

This is a very basic point with respect to many new sources of travel data. For example, mobile phones log where a call is made from and to, and these can be extracted from the call detail record (CDR). With judicious analysis of times when such calls are made and the relative longevity of time spent in fixed places, judicious estimates of the correlation between physical travel patterns and the flow of calls can be made. In fact so far although these data look promising, they are very hard to integrate and improve without independent data that can be related to them. Doing the same with Twitter and other social media data is equally problematic. The key challenge is shown in Figure 19.2 where we have data on origins and destinations and trip patterns from small areas in Greater London for travellers who use a variety of tube, road, bus, rail, bike and walk modes to get to work (Figure 19.2a, c, b) and

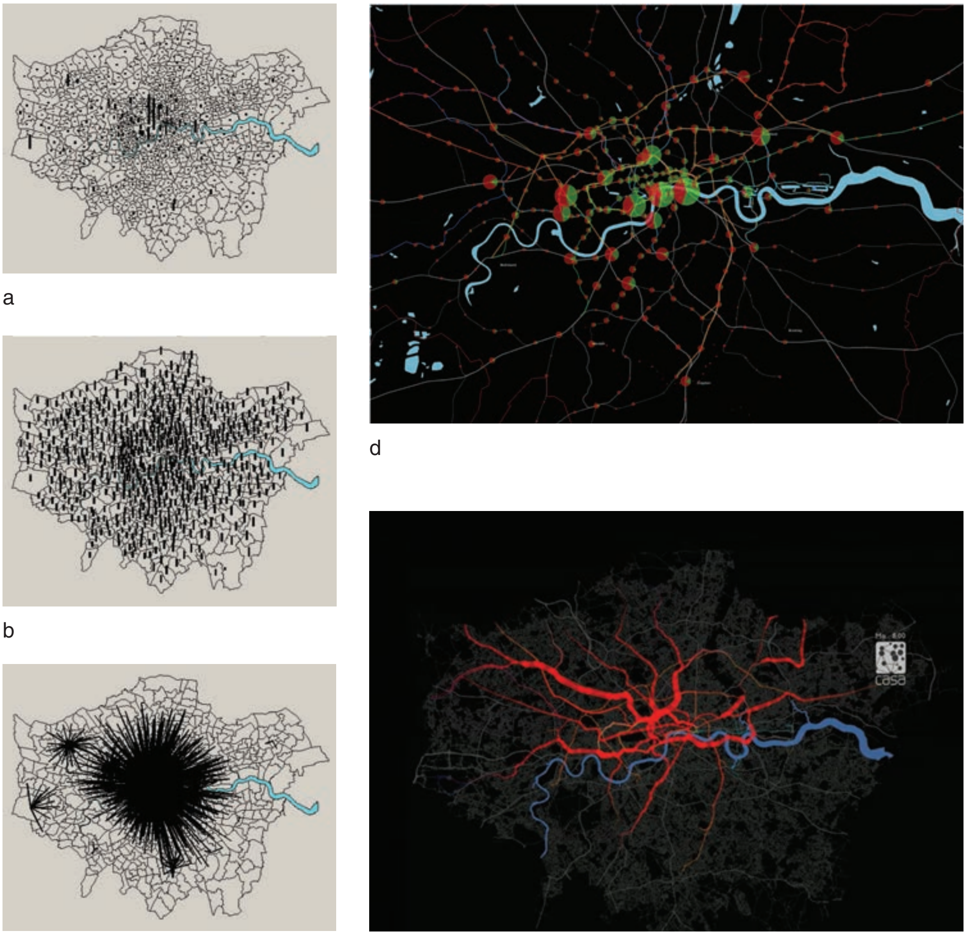

FIGURE 19.2 Traditional transport flows from the population census by wards compared to tap-ins and tap-outs from Transport for London's Oyster Card data. (a) Origin employment (2001) by wards, (b) destination working population (2001) by wards, (c) trips (2001) greater than 200 persons per ward to ward flow, (d) tap-ins and tap-outs at 8:00am on the London Tube (November weekday 2010), (e) flows of tap-ins and outs assigned to the tube network.

the flow of data on the tube at the peak hour where we assume that travellers take the shortest routes between the points where they tap-in and tap-out using their smart cards (from the Oyster card data in Figure 19.2e). These two sets of patterns might be supplemented by other detailed flows by mode and volumes of traffic on the highway from loop counters. The picture we are able to get is highly variegated with respect to error and ambiguity and as yet we do not even know if these sorts of data coming from real-time streaming in the smart city are going to be useful for the traditional practices of transport planning.

These transport flows are material movements in that they are measured, here in persons but they could be commodities, freight and so on. The smart city is also being fast automated with respect to flows of information measured in terms of communications using social media, financial transactions and email–web related traffic. There are many diverse sources of such data and to illustrate one, the Spanish BBVA Bank have produced data from their credit card transactions which have been animated to show spending patterns over short periods of time. The movie produced by the Senseable City Lab at MIT shows such flows over Easter 2011[1] and Figure 19.3 shows two infographics from data that pertain to financial flows that are being used to figure out where people shop and how much money they spend. Figure 19.3a, b and c shows origin and destination maps of sales and a

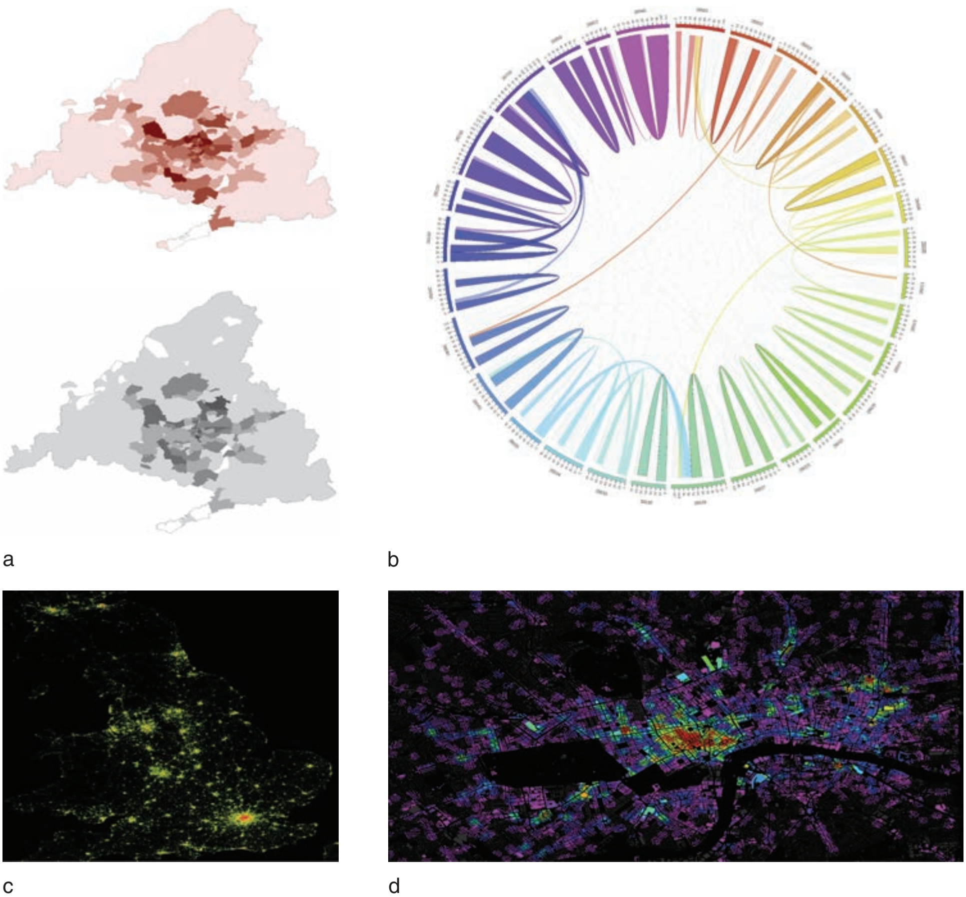

FIGURE 19.3 Credit card data for retail sales in Madrid and Foursquare check-ins from the social media correlated with retail sales in the UK and London. (a) Origins and destinations of sales in Madrid with (b) a circular visualisation using Circos software of flows that show most inside the relevant zonal area, (c) density of Foursquare check-ins in the UK, (d) similar at the level of retail locations in Central London.

circular flow graph of the flows of credit card transactions aggregated to 32 locations in Madrid. Figure 19.3c and d shows the clustering of Foursquare check-ins defining shopping patterns in the UK and Central London, which we can be used as a proxy for where people shop.

CROWDSOURCING, OPEN DATA AND VALUE THROUGH ANALYTICS

So far it is clear that the exhaust from real-time control, which is one of the missions of the smart city movement, provides many new data sources that inform us faster and more precisely than ever before, but also raise important problems of representativeness. In fact, it is fairly obvious but needs reinforcing that it is the World Wide Web – the internet or net of networks – which is the new glue that is tying us and our objects and artefacts all together. It is an internet of things that is driving the new ways in which we can share data, simulations and forecasts amongst ourselves with the prospects that all of us can engage in altering those very data, models and forecasts that in the past remained the prerogative of the very few. New ways of opening data are now possible, but only because of the web. So much government data are now in the public domain, such data having been deregulated, and this is enabling many new software products to be developed that add value to such data (see Chapter 9).

Open data, however, is simply the tip of the iceberg with respect to new and big data that are being sensed, collected and collated. Data that are collected by ourselves through various forms of crowdsourcing – mobilising the power of the crowd who invariably are responding to mandates to provide data and opinions in some semi-structured fashion – are becoming ever more significant. Much of the data sensed in this way through online questionnaires, are providing data beyond that which are sensed in real time. Much of these data are controversial as they reflect interpretations and opinions that cannot often be laid bare, but some are quite objective. OpenStreetMap (OSM), which has been built worldwide by a generation of those interested in location and place and tagging using global positioning systems (GPS) on smartphones is an excellent case in point. Much of the data in volume terms are as good as those collected by national mapping agencies. There are many such examples of crowdsourcing, but for it to be really effective, it needs a broadcast medium to tell the crowd that there are those who want to elicit their opinions. Some of this is done by the internet itself and by the use of social media, but the best is where there is a medium like TV that is widely accessed and through which the crowd can be mobilised. An example that our group has worked upon was based on local TV media running a campaign to elicit responses to key problems of social behaviour and to enable TV watchers to log and record their responses to the relative importance of social problems by recording their responses and their locations. Work with the BBC local news programme *Look East* elicited many thousands of responses across the eastern region of England concerning the preponderance of key problems involving drunkenness, stolen cars, noise, pollution and so on (Hudson-Smith et al., 2009). In Figure 19.4, we show an example of this with a

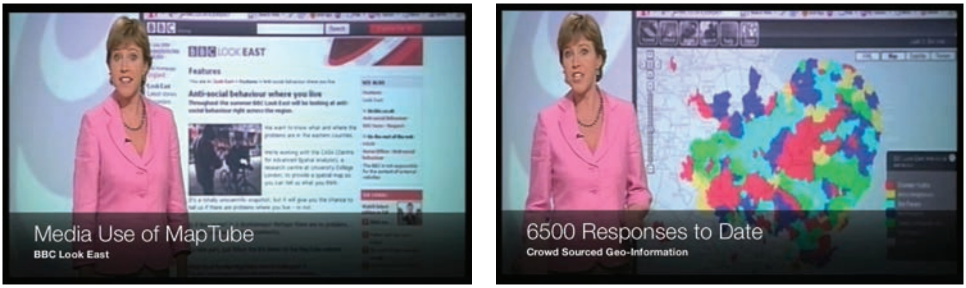

FIGURE 19.4 Crowdsourcing: a broadcast medium to reach TV viewers

shot from the broadcast TV medium alongside the map that was produced over a short period of time, it being continually updated as more and more watchers responded.

There are many different variants on this notion of collecting data from the crowd, which range from routine lower level responses, to relatively unambiguous queries, to scientific sourcing where users are enticed to solve problems, discover new ideas and present their insights into key controversies in a structured way. One of the things that has not been introduced so far is any review of the range of theories and analytical techniques that are being fashioned to deal with these kinds of data. New statistical models are clearly needed as noted with respect to dealing with datasets whose structure and volume are qualitatively different from previous data volumes that were much smaller. New methods of data mining are important now that are able to extract deep structures and structures that vary widely across any single dataset. In short, the bigger the dataset, the more variegated the structures that are possible within it, and our models and techniques need to be adapted to take all this into account. Respondents are asked to answer to questions about social behaviour in their own locality and to record such responses using CASA's online mapping system MapTube.[2]

You might anticipate that a new science of cities is developing by building on what has gone before where an older science produced analytics well adapted to our prior understanding of cities and the needs in their planning. Nothing could be further from the truth. Cities, like all human systems, are continually in flux and their complexity appears to be outstripping our abilities to make sense of them. The previous section on transport flows provided a glimpse at the past generation of models that deal with flows, but the new data which are streamed in real time change the focus quite radically and as yet we have not been able to integrate models that deal with different temporal resolutions and intervals. New models and new forms of analytics that go beyond data mining are urgently needed and although there is some prospect that more disaggregate models – agent-based models – might hold the key to integrating different spatial and temporal scales, the field is wide open for new insights and new theories that make sense not only of the smart city, but of the wider context of spatial media of which the smart city is a part.

CHALLENGES FOR STAKEHOLDERS AND CITIZENS

A final feature of the smart city relates to those who live in it – ourselves, acting as citizens and stakeholders. The web is bringing a wide variety of services to us that we have traditionally travelled to consume or at least collect as well as services that are entirely new and a product simply of the fact that we can now access them using IT. Every piece of information that traditionally required us to manually complete paperwork can now be delivered by the internet and processed online. In principle, all bureaucratic transactions can now be completely recorded, but the data volumes are huge and as in earlier times when these tasks were not automated, organising and understanding their wider import is as significant as it ever was. Elsewhere in this book, there is discussion of the extent to which this new media is creating a citizen science (see Chapter 12), new forms of participation (see Chapter 18), questions of privacy and confidentiality in data (see Chapter 22), as well as the limitations and potential of all these new ways of working. These are all part and parcel of the smart city.

The challenges we face in the development of smart cities are enormous. Large IT companies who see the smart city as just the next phase of the roll out of their business have found a new mission in wrapping up traditional ways of making money, seemingly invoking a mission with social objectives. Much of the smart city rhetoric is phrased in terms of achieving a better quality of life for citizens, but the debate is strangely silent about questions of segregation, inequality and poverty, and tends to focus more on accessibility and economic opportunity. This is where the wider constituency of academics, policy-makers and citizens at large have a critical role to play. In leading the debate about how new information technologies can be adapted, ways in which both public as well as private goals that the community and its individuals might achieve in the wider context of such automation can be articulated and reached.

ACKNOWLEDGEMENTS

Ollie O'Brien devised the original London dashboard and Steven Gray extended the media both within it and the way it displays. Jon Reades was instrumental in developing our work with Transport for London's Oyster Card data; Joan Serras for the work with the BBVA credit card data for retailing in our EUNOIA project; and Vassilis Zachariadis for the retail model that is using Twitter and Foursquare data to build proxies for retail attractions in our INSIGHT project. Richard Milton devised MapTube and was central to the Look East crowdsourcing project.

NOTES

1. https://www.youtube.com/watch?t=13&v=8J3T3UjHbrE
2. http://www.maptube.org/

REFERENCES

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

Batty, M. (2016) 'Cities in disequilibrium', in J. Johnson, A. Nowak, P. Ormerod, B. Rosewell, and Y.-C. Zhang (eds), *Non-Equilibrium Social Science*. Berlin: Springer.

Batty, M. (2015) 'The sixth Kondratieff is the age of the smart city', Spatial Complexity. Available at: http://www.spatialcomplexity.info/files/2015/05/Draft-Chapter-2015.pdf (accessed 27 July 2016).

Grech, G. (2015) 'Cities as platforms', *Tech Crunch*, 7 August. Available at: http://techcrunch.com/2015/08/07/cities-as-platforms/ (accessed 27 July 2016).

Hudson-Smith, A., Batty, M., Crooks, A. and Milton, R. (2009) 'Mapping for the masses: accessing Web 2.0 through crowdsourcing', *Social Science Computer Review*, 27 (4): 524–38.

Kitchin, R., Lauriault, T.P. and McArdle, G. (2015) 'Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards', *Regional Studies, Regional Science*, 2: 6–28.

Meier, R.L. (1962) *A Communications Theory of Urban Growth*. Cambridge, MA: The MIT Press.

O'Brien, O., Batty, M., Gray, S., Cheshire, J. and Hudson-Smith, A. (2014) 'On city dashboards and data stores', a paper presented to the Workshop on Big Data and Urban Informatics, 11–12 August 2014. Chicago, IL: University of Illinois. Available at: http://urbanbigdata.uic.edu/proceedings/ (accessed 27 July 2016).

Webber, M.M. (1964) 'The urban place and the nonplace urban realm', in M.M. Webber, J.W. Dyckman, D.L. Foley, A.Z. Guttenberg, W.L.C. Wheaton and C. Bauer Wurster (eds), *Explorations into Urban Structure*. Philadelphia, PA: University of Pennsylvania Press. pp. 79–153.