Smart cities, big data and urban policy: Towards urban analytics for the long run

Jens Kandt*, Michael Batty

The Bartlett Centre for Advanced Spatial Analysis (CASA), University College London (UCL), Gower Street, London WC1E 6BT, UK

ABSTRACT

The analysis of big data is deemed to define a new era in urban research, planning and policy. Real-time data mining and pattern detection in high-frequency data can now be carried out at a large scale. Novel analytical practices promise smoother decision-making as part of a more evidence-based and smarter urbanism, while critical voices highlight the dangers and pitfalls of instrumental, data-driven city making to urban governance. Less attention has been devoted to identifying the practical conditions under which big data can realistically contribute to addressing urban policy problems. In this paper, we discuss the value and limitations of big data for long-term urban policy and planning. We first develop a theoretical perspective on urban analytics as a practice that is part of a new smart urbanism. We identify the particular tension of opposed temporalities of high-frequency data and the long durée of structural challenges facing cities. Drawing on empirical studies using big urban data, we highlight epistemological and practical challenges that arise from the analysis of high-frequency data for strategic purposesand formulate propositions on the ways in which urban analytics can inform long-term urban policy.

1. Introduction

Big data and, more generally, digital technologies are regarded as paramount in the governance and planning of smart cities. Numerous scholars in urban research argue that a new type of big data analytics promises benefits in terms of real-time prediction, adaptation, higher energy efficiency, higher quality of life and greater ease of movement (Batty, 2019; Batty et al., 2012; Kourtit et al., 2017; Townsend, 2013). Ubiquitous digital technology embedded in the physical structure of the city makes these desired outcomes possible. The vision is that assemblages of sensing devices linked through the Internet of Things to computing platforms process incessant streams of data and thus enable new ways of decision-making, which, in some contexts, can be automated with little human intervention (Harrison et al., 2010; Rathore et al., 2016).

Urban analytics, which we define as big data analytics applied for the purpose of urban governance and planning (Kang et al., 2019), is held to be central to shaping and running smart cities. The belief is that, as the availability of real-time data and the necessary computing power continue to increase, so does the possibility of detecting patterns in the everyday life of urban systems and developing applications that respond to these patterns in real-time (Batty et al., 2012). Among the various applications in urban transport, energy or policing, intelligent

traffic management is perhaps most established. New systems, such as Siemens's Real Time Optimiser currently developed for London (TfL, 2018), employ analytics to predict traffic congestion and enable efficient use of road space in real-time. The system is also expected to achieve strategic policy goals including increased road safety, improved air quality and public health. The assumption is that analytics of urban high frequency data can contribute to the solution of long-standing challenges in cities.

Yet, widely articulated visions on the strategic value of big data for cities bridge fundamentally different temporal scales of urban dynamics: the short-term scale of fast dynamics or real time and the long-term, much slower dynamics of urban structure and policy. This tension, which arises in part out of the contradictory 'timescape' of smart cities (Kitchin, 2019), warrants closer examinations with regard to the practical conditions under which big data and urban analytics can productively contribute to strategic urban policy and planning.

The contrasting temporalities of high-frequency data and real-time response on the one hand and the long durée of structural urban dynamics on the other hand generate several epistemological challenges, some of which can be associated with positive feedbacks or unintended consequences. To take our previous example, analytics-powered, intelligent traffic management may make day-to-day travel smoother and thus in turn encourage further travel - a phenomenon known as

---

'induced demand' in transport planning. Beyond travel choices, smoother travel may encourage residents to move and take advantage of lower housing cost further away from their workplace, thus influencing the diurnal and residential geography of the city. Such changes, though they are triggered by momentous and, in this case, algorithmic responses, unfold over very different time scales and in turn alter those high-frequency data streams affecting system responses, whose consequences and socio-spatial implications are hard if not impossible to predict (Batty, 2018).

Such epistemological and strategic challenges call for a critical discussion of how the fast temporal scale of big data can be reconciled with the slow dynamics of the socio-spatial organisation of cities. In particular, the different temporalities compound the problem of causality, that is how observations in high-frequency data can inform the causal insights needed in urban policy and planning. Being mindful of powerful social science critiques of urban analytics with their renewed concerns about technocratic, data-driven approaches to urban governance (Kitchin, 2014a; Rabari & Storper, 2015; Wyly, 2014), we wish to join efforts in critically gauging the value of big data (Kitchin, 2016a; Singleton & Arribas-Bel, 2019) by specifically focusing on issues of temporality and causality in urban systems (Batty, 2018; Kitchin, 2019). Our aim is to highlight the conditions under which urban analytics of high frequency data can realistically and productively inform long-term urban policy.

In the following, we first summarise current debates on and critiques of data-driven smart urbanism and urban analytics. We then draw on contemporary social science theory to chart the social genesis of urban analytics and conceptualise its practices in the context of hypermobile societies. Subsequently, we highlight salient characteristics of urban analytics and identify practical, epistemological issues arising from the analysis of fast urban dynamics for long-term urban policy. Finally, we develop six propositions to identify the conditions under which urban analytics can productively inform strategic, long-term urban policy. We conclude with suggestions for research investigating and extending these propositions.

2. Smart cities and urban analytics

The term urban analytics has arisen in the context of processing big data for a range of applications in cities, in particular 'smart' cities (Batty, 2019). Smart cities are variously characterised as the emerging, future version of the city that runs in part on streams of data continuously flowing between physical objects, actors and institutions that define, inhabit and govern cities (Batty, 2020; Batty et al., 2012; Townsend, 2013). Smart cities are collections of numerous sentient and connected built environments, which possess components that learn from patterns of daily activity and adapt automatically to changes in such behaviours. This is made possible by the scaling down of computers to the point where they can be embedded in the everyday objects and activities alongside advances in sensing and computing power, all of which allow for the ubiquitous integration of microelectronics into the physical world. To employ the physico-digital infrastructure that collects and processes data and integrates them in decision-making, cities and private sector companies regularly team up together in public-private partnerships to carry out the necessary large-scale investments (Hollands, 2008; Kitchin, 2014b).

Smart city initiatives have now been subject to extensive academic debates, which range from attempts to develop formal definitions to radical, social-science critiques. Existing notions and perceptions of urban analytics are closely linked to those debates, which, to a degree, can be associated with different epistemological viewpoints that prevail in the social sciences. One strand in the literature conceives of the smart city as a distinct, definable and specifiable thing that is or will become physically and institutionally manifest, thus knowable and measurable (Albino et al., 2015; Komninos, 2011; Komninos et al., 2013; Neirotti et al., 2014). Conforming to positivistic conceptualisations of the social

world, this literature views the smart city as a value-free, technological solution built on objective evidence on urban activities and inevitable human progress. Smart cities are presumed to embody a consensus on how cities should be organised, designed and managed and thus as such they represent a desirable urban vision. Consequently, research and development should be devoted to specifying appropriate system architectures, management frameworks and universal technical standards (Abella et al., 2017; Babar & Arif, 2017; Rathore et al., 2016; de Santana et al., 2019; Zygiaris, 2013). In this context, urban analytics is a set of technical methods, which are applied to the digital infrastructure and include the development of algorithms. The methods are scientific in character; they ensure optimal data processing to enable intelligent, efficient and automated responses while affording factual insights into urban systems to decision-makers.

A second strand in the literature that views smart cities as deeply political projects designed to surveil and control urban citizens and to extract profit from such activities. The most radical critiques portray smart cities as hegemonic projects within the neoliberal political economy, in which governments and corporations ally to employ digital technology to consolidate and extend their power over all domains of society across the globe (Greenfield, 2013, 2017; Hollands, 2015; see also Zuboff, 2019). Softer, critical accounts identify networks of corporations, policy advocates and technocrats as particularly powerful actors who persuade often under-resourced city governments to adopt wholesale smart city solutions (Barns et al., 2017; Gaffney & Robertson, 2016; Söderström et al., 2014). Examples include the corporate-led urban projects of Songdo or Masdar or supposedly off-the-shelf infrastructure products such as urban 'operating systems', control centres or data stores in Rio de Janeiro or New York. Even in more state-regulated societies, such as Singapore, large IT companies are leading the drive to automate and monitor people's movements associated with e.g. more efficient use of various kinds of transit. Urban analytics is viewed as a core ally in the 'datafication' and 'dataveillance' that is central to such smart city projects (Kitchin, 2016b). Here, urban analytics is a political tool of control through 'software-sorting' (Graham, 2005), profiling people and places according to their characteristics and potentials in terms of risk or propensity to conform. It is deliberately presented as data-driven, value-free and pragmatic but really designed to reproduce if not amplify prevalent power relations (cf. Kitchin, 2014b; Söderström et al., 2014; Vanolo, 2014). At best, urban analytics is employed naively, lacks transparency and suffers from unacknowledged social bias, a view that resonates with more general critiques of positivism and its politics.

In a third perspective, smart cities are principally imaginaries of future cities that drive contemporary approaches to city making. Such imaginaries are produced, articulated and mobilised by agents situated in multiple, social networks, of which international networks have gained particular significance (Kitchin et al., 2017; Luque-Ayala & Marvin, 2015). Participants in these networks of corporations, experts, advisers, think tanks and scientists act in contexts in which certain beliefs, ideologies and practices prevail. Through diverse practices involving visits, face-to-face meetings, policy papers and smart city benchmarking exercises (Bok & Coe, 2017; Campbell, 2012; Prince, 2015), smart urbanism becomes manifest as a new way of planning and governing cities, a perceived imperative of technological progress to confront the challenges in and of the 21st century. Yet there is no guarantee that visions will be implemented in perhaps the hegemonic way in which some powerful players would like to see them (Kong & Woods, 2018; Wiig, 2015). For example, city governments in Europe, being situated actors, often seek to balance demands arising from smart city visions against the content of their political mandates (Bunders & Varró, 2019), although they are regularly receptive to smart city solutions as part of their wider entrepreneurial orientations (Hollands, 2008; Vanolo, 2014). Projects may therefore be contested, redefined, modified, adapted or even terminated, thus exhibiting the well documented dynamics of policy mobility (Peck & Theodore, 2010).

In all these perspectives, it is clear that there are influential groups and networks that promote a smart urbanism in which processing of high frequency data is implicated in various ways. We therefore suggest that urban analytics comprises a set of practices that are themselves part of a bundle of practices that constitute smart urbanism, and in parts, a data-driven form of city-making (Boeing et al., 2020). The growing attention to data science in city governments attests to the proliferation of beliefs and convictions about the success and desirability of quantitative prediction, which align with positivistic and technocratic visions of society (Delmelle, 2019; Kitchin et al., 2015; Wyly, 2014). Urban analytics practitioners tend to share beliefs, ideologies, unconscious subjective orientations and dispositions - professional or otherwise - that prevail in networks promoting smart urbanism; but they are also increasingly located in certain parts of academia, exemplified by the rise of 'computational social science' (gesis.org/en/institute/departments/computational-social-science), and marked by increasingly distinct scientific practices focused upon data mining, pattern detection, machine learning and prediction.

Therefore, we can say that, rather than a set of clearly defined technical methods, urban analytics comprises a set of practices that are carried out in specific social, political and academic contexts. This 'practice notion' of urban analytics can be related to a wider body of theories of 'social practices' as conceived by sociologists (Bourdieu, 1977; Giddens, 1984) and reworked in the context of 21st century mobilities (Sheller, 2017; Urry, 2007). Before we examine urban analytics as a practice and its possible role in urban policy more closely, we chart the societal trends that give rise to the wide-spread embedding of sensing and processing technologies in urban environments.

3. Mobilities and big urban data

Among the long-standing debates concerned with the rise of digital technology and its implications for cities (Batty, 1997; Castells, 1996; Graham & Marvin, 1996, 2001; Larsen et al., 2006; Lash, 2002; Thrift & French, 2002; Urry, 2000, 2007), the Mobilities paradigm offers a recent, practice-based perspective through which the embedding of digital technology in cities can be conceptualised. The core tenets are that mobilities define social relations in the 21st century (Sheller & Urry, 2006; Urry, 2007, 2012) as the dialectically linked co-evolution of Transportation, Information and Communication Technology (which we refer to as TICT henceforth) and socio-spatial relations that have resulted in social interactions that are increasingly far flung. Social relations are now formed and sustained over larger distances through travel as well as new and vast amounts of electronic communication, including virtual face-to-face contact. Despite the existence of tele and electronic communication, strong social links marked by trust still require direct face-to-face contact and co-presence (cf. Giddens, 1984, 86; Urry, 2007, 166). Notwithstanding the possibility that this sociological law might change in the future, it implies that, as social agents form many more weak links through increased travel and virtual presence, further, physical travel is necessary to advance and sustain these social links.

As a result, social relations and, to a degree, the functioning of society requires travel and digital communication over increasing distances, which depend on TICT bundled up in increasingly complex infrastructure, or what Urry (2007) calls 'mobility-systems'. Such systems include railway networks, the car, mobile phone networks, airports, in fact one could add entire multi-modal systems if not even the smart city itself in its entirety, which in many senses is all about enabling mobility and organising flows of people, goods and information. In so doing, any mobility-system becomes indispensable in sustaining the communications and flows that are crucial to the functioning of society, and in a self-reinforcing loop, generate further far flung social interactions, which demand ever more travel and communications.

In sustaining and expanding the space-time distanciation of social relations, mobility systems have become increasingly digital in the

sense that digital technologies have become more widely embedded in the physical environment. As social interactions over larger distances result from digitally infused and controlled TICT, which in turn drive a barely reversible process of further digitalisation and growing reliance on expert forms of knowledge (see also Kitchin & Dodge, 2011, 10), we may argue that, in processing the data generated by physico-digital mobility systems, urban analytics is a practice that forms part of the expanding mobilities that characterise 21st century society. Increased computational power arising out of the same social and economic forces fuel the new practice of urban analytics and engender the new analytical, big data procedures of real-time processing, pattern detection and statistical learning used to identify and execute optimal, automated responses.

Mobilities thus offer a social explanation for more enthusiastic visions and advocacy of smart cities, such as proposed data-driven, integrated urban management and planning (Babar & Arif, 2017; Rathore et al., 2016; de Santana et al., 2019) and the recently enacted ISO certified smart city standards, which specify everything from data interoperability to urban governance arrangements (BS ISO 37106:2018, 2018). We can also observe a trend in which movement is increasingly tied to digital identities linking up various digital traces of individuals, be they biometrics, GPS locations, credit card purchases or social media activity. Recent examples include the Known Traveller Programme (WEF, 2018) or Social Credit Systems (Merics, 2017), which can be used to 'software-sort' and regulate the behaviour of individuals, organisations or institutions (Graham, 2005; Wood & Graham, 2006).

In the same vein, the emergence of the computational social sciences can be viewed as a result of contemporary mobilities. This discipline, characterised by frequent cross-sector collaborations between academia and business, are supported by remote processing and cloud computing, which further mobilities-related technologies enabling collaboration and joint analysis over large distances. Such novel practices occur against a wider backdrop of a mobilities-dependent academia, which, much like any other profession, sustains access to networks, knowledge and other resources through electronic communication, online publications and, until recently, increased international travel (Storme et al., 2017).

Finally, much like mobilities, we can expect that urban analytics develop self-perpetuating qualities, which go beyond the satisfaction of specific needs. Circulation can be a source for profit, perhaps best exemplified by the car industry. The potential profits arising from providing the technology for processing big data reinforces urban analytics, its endorsements and associated practices. A few examples include an actively voiced enthusiasm about the detail and volume offered by new, big data sources, an appreciation of efficient, optimised urban operations, an interest in the fast dynamics of the city and real-time prediction, and the promotion of data-driven decision-making (see for example, Provost & Fawcett, 2013). We can expect that irrespective of its potential contribution to social goals, urban analytics is actively promoted by the same 'epistemic communities' and 'advocacy coalitions' (Kitchin et al., 2017) that mobilise smart urbanism in a wide range of contexts.

It is clear then that mobility systems and their physical and digital components have a profound impact on practices and relationships between actors of all kinds. Kitchin (2019) describes how digital and mobile technologies alter the pace and rhythms of daily lives of both citizens and city operations. 'Realtimeness' (Kitchin, 2019, 786), enabled by big data, its underpinning digital and networked technologies and the practice of urban analytics that they make possible, produce this temporality that characterises smart urbanism. Kitchin perceives a tendency in city operations to privilege real-time response over planning, reflection and long-term strategizing. This tendency may be further amplified by an awareness of the amount of financial commitment that is necessary to sustain 'realtimeness' across areas and sectors in the city.

For the relationship between urban analytics and policy, the

question of temporality poses two epistemological and thus practical problems. First, there is a need to examine the conditions under which high frequency data can help us understand cities beyond momentary, fast, perhaps transient phenomena. In other words, what kind of 'understanding' of the city is promoted by the analysis of fast dynamics, and to which degree do insights apply to general and deep-seated causal dynamics urban policy is typically concerned with?

A second problem relates to the fact that urban analytics uses the same technologies that algorithmically alter, synchronise or regulate the practices of urban actors to study and understand cities (Batty, 2019). Transit smart cards both capture but also alter travel demand through the automated interventions the systems are designed for (e.g. dynamic peak/off-peak pricing of fares), which affect our ability to understand people's mobility through their data. As a consequence, any patterns we observe are as much a product of human subjects as of the workings of the technologies themselves; their genesis is rooted in the human and institutional dynamics of mobilities of contemporary society. This fundamental question of causality alongside the bias inherent in big data (Boyd & Crawford, 2012; Taylor, 2015) are critical issues for long-term urban policy and planning, which, with their objective to organise and improve cities, essentially seek to intervene into complex, causal relationships in urban systems.

4. Making sense of urban analytics

To identify the salient characteristics of urban analytics as a practice, we now review a number of urban analytics research applications that come from a wider array of real time analysis of mobilities using travel smart card data. Travel smartcards automatically collect fares on public transport systems and thus produce large volumes of travel records pertaining to the vast majority of passengers. The advantage compared to traditional data sources such as surveys of passengers are their coverage, large sample size and precision. Although smartcard schemes have arisen out of the need to facilitate and refashion transit fare collection to make daily travel smoother, the data source is also widely viewed as a strategic asset for transport planning (Anda et al., 2017; Pelletier et al., 2011). While the vast majority of studies process smartcard data for transport modelling purposes, some studies seek to characterise travel demand in novel ways drawing on the detail such data offer. Only a very small number extend the analysis beyond a few months to a longer time frame.

4.1. Characterising fast urban dynamics in daily movement.

Numerous studies demonstrate how the coverage and precision of smart card records enable novel characterisations of travel demand (e.g. Goulet Langlois et al., 2016; Liu et al., 2019; Reades et al., 2016; Zhong et al., 2014; Zhong et al., 2016). The analytical approaches to process smart card records are diverse but often contain a crucial step of exploratory data mining to extract patterns. Many studies focus on developing novel ways to characterise urban mobility and limit their analysis to one snapshot in time.

To demonstrate the extent to which the volume of this type of high frequency data set can inform policy applications, we turn to our own analysis of Oyster card data in London. Oyster card data are now used to pay for 85% of the journeys taken on the London Underground (the Tube). During a typical weekday, up to 5 million passenger journeys are made on the tube, there are 11 underground lines and 270 stations. We use the journeys extracted from the card which lets a passenger tap into a station at the start of the trip and ensures that they tap out at the other end, otherwise incurring a penalty. The data have no attributes other than the time and place of the tap-in and tap-out, the status of the card with respect to payment, and whether or not a passenger is a senior, child or disabled person.

This data set is comprehensive and detailed with respect to the origin (O) and destination (D) of the trip (the stations), not the route on which the passenger travels. We infer the routes by determining the shortest route from any one station to any other using a variant Dijkstra's (1959) standard algorithm and then extracting the origins and destinations for each trip and tagging this with the shortest path. The distribution of these paths in and of itself provides detailed and novel information as to the redundancy in all possible routings within the transit system. But such trajectories also help to capture and characterise the heterogeneity of travel habits at the level of minutes. Various studies suggest that daily activities of travellers can be inferred from the space-time characteristics of repeated trips over a given time period (Reades et al., 2016; Zhong et al., 2014, 2016).

Examining any profile of trips for any station over the shortest time period we consider analytically acceptable such as 1, 2, 5, 10 min, we find that the temporal incidences of individual trips differ on each day of the week. This is because passengers cannot arrive at exactly the same time each day at a station but it is also because there is great heterogeneity in the profile of each traveller with respect to the working week. From the data we can extract how many people have similar profiles with respect to any particular station and the time of entering or exiting. We find that only about one third of passengers use the same tube stations for exit and entry twice a day, which would imply the kind of regularity associated with a journey to work or some regular daily visit associated with two different places. Although the morning and evening peaks and the late-night evening peak in the centre of London are clearly marked routines, their composition day-to-day is quite varied implying that more than half of all passengers do not have regular routines. Such findings remain masked in aggregate and low sample data; they can only be identified through mining high-frequency data at the scale of individual records.

To explore regularities in disaggregate patterns, we aggregate across stations and profile them with respect to similarities between any and every pair of trajectories. We choose any two time periods and compute the similarities between the volumes of trips at each station – the number of taps in and out – with respect to how they compare across all 270 stations. In short, this comparison is between any two times (from 72 in total) with respect to the volumes of trips at each of these time periods across all 270 stations. Time is measured in 20-minute intervals over the 24-hour data, thus giving 72 time periods. We chose this interval as it is both sufficiently granular to highlight in sufficient detail how different any one time period is from any other across the 24 h of the analysis. Of course, big data allow flexible granularities any finer time bins can be chosen. The 20-minute time bins allow us to examine the patterns visually and guide decision on further analysis.

The profiles with respect to any pair of time periods are shown in Fig. 1 where the first map (A) shows the correlation density between any pair of time periods with respect to the volume of trips across all stations; that is the correlation is a suitably normalised function based on comparing the variations of the flows in all stations over any two time intervals. We have also computed the similarity between any two trajectories with respect to the ranks not the volumes of trips. Here we rearrange the profiles according to the rank of volume and then make the comparison as a normalised difference between each pair of the ranked profiles over stations in each time period. We show the pattern in Fig. 1(B) while in Fig. 1(C) we compute the same set of similarities but this time in terms of the absolute volumes. This figure is closer to Fig. 1(A) for the measure used simply smooths the patterns while in Fig. 1(B), the structure of the travel profiles is sharpened.

It is clear from Fig. 1(A) that the morning peak – the first set of rows with high correlations between 7 and about 9–20 in the morning imply that these are fairly similar trip distributions. This means the volumes across all stations correlate with each quite highly in the 20-minute intervals that define this peak. Then from about 10 am to 3.30 pm there is a similar block, which can be interpreted as being a different sequence of profiles for the hours in the middle of the day. There are a lot fewer changes in total trips then and a lot less overall number but the matrix still brings out the similarities in the temporal patterns not the

total volumes. Then from about 4 pm to 7 pm there is the evening peak and this blurs both into the middle of day and the mid-to-later evening volumes. When we sharpen the data by rank ordering everything and computing a similarity coefficient between the ranks as in Fig. 1(B), the same pattern emerges but this is much clearer. The middle of the day and the early evening peak merge into one another while the block of evening travel profiles are quite distinct. Fig. 1(C) is something of a halfway house which simply differs from Fig. 1(A) due to a different measure being applied to the same data profile.

To provide detailed interpretations of the travel profiles, we note that one needs to make informed assumptions about the nature of urban travel and the place in question, including the structure of trip distribution in western cities during the working week and the associated peaks in demand at various times indicating different urban activities. In the case of data where we have to extract the structure, then even the ways of visualising the data, including the size of time bins, need to be known before analysis begins. Therefore, visualising, recognising and interpreting patterns is a circular process that relies on external data, information and theory about how the urban system functions. This reliance increases in analyses of high-frequency data for long-term applications.

4.2. High frequency data for longer time frames

Studies that process smart card data over time frames that are longer than a few months are rare. Chu (2015) analyses smartcard records spanning two years and to extract travel trends at individual and aggregate levels. Briand et al. (2017) link all transactions recorded in February across five years to characterise year-on-year change in bus use. Their segmentation-based data mining approach reveals a complex assemblage of changing temporal travel patterns masked within an overall trend of stable bus patronage. Huang et al. (2018) process a small sample of smartcard transaction over seven years to investigate interdependencies between residential mobility, job change and daily travel. The emphasis of this small number of pioneering studies lies on characterising trends at different levels rather than systematic attempts to explain them.

The question of causality is central in our own research exemplified by Kandt and Leak (2019), who process six years of smartcard data in the metropolitan region of Birmingham, UK. The region's transport authority, Transport for West Midlands, has recorded a steady decline in bus patronage in particular among older residents, who are entitled to travel for free. The decline in public transport patronage stand in direct opposition to the city's long-term goal of ensuring environmental sustainability, social inclusion and health in parts through ensuring equitable access to public transport.

Smartcards were introduced in 2010 to record all concessionary journeys for the purpose of reimbursing operators. Sequence analysis is an unsupervised data mining technique that can detect profiles of similar temporal trends in in longitudinal data (Gauthier et al., 2014). Fig. 1 shows an extraction of six temporal profiles found in 300 million smartcard transactions made by passengers aged 66 or older between 2011 and 2016 (source: Kandt & Leak, 2019). Two of the groups indicate a pronounced decline in boardings during the six years.

In order to interpret the patterns in terms of causality, we drew on low-frequency contextual data, the UK Census neighbourhood statistics and official mid-year population estimates. The contextualisation surprisingly revealed that a particularly sharp decline in patronage occurred in poorer neighbourhoods, where dependence on free bus travel is usually higher. Based on the contextualisation and some considerations mainly drawn from the travel demand literature, we were able to develop policy-relevant causal hypotheses (Table 1). This process led to a final step – triangulation, gauging the plausibility of hypotheses, using another set of small and low-frequency data as evidence, including reports on online shopping trends and pensioner poverty and testimonial pieces of evidence, such as the time when and the way in which e-

hailing services where permitted in the region.

The study shows that the derivation of any potential causal insights from high-frequency data for long-term policy requires a number of elements: theory-informed interpretation, contextualisation and triangulation. This contrasts with a purely computational approach, which would consist in taking the patterns per se and designing applications prompting passengers matching a particular profile to change their choices in order to help reverse the trend. Such an automated approach, often envisaged as part of smart urbanism, may be effective in achieving short-term goals, such as increasing revenue through better targeted marketing (Singleton & Arribas-Bel, 2019, 10). But in an urban planning and policy context, this would equate to treating symptoms and not the structural or social causes of issues related to inclusive mobility. Such social causes may find expression in sensed data, but they cannot be read off from these emergent patterns themselves.

How then do we interpret patterns in big data? While we are not aware of any empirical work investigating the cognitive processes involved in interpreting big data, social practice theory and its supporting evidence suggests that interpretations will inherently be influenced by the beliefs, day-to-day experience and subjective orientations of the interpreter and those beliefs and dispositions that prevail in the networks the interpreter is situated in (see also Boyd & Crawford, 2012; Taylor, 2016). The influence and elusiveness of tacit interpretation constitutes a fundamental difference to the processing of conventional data sources, such as purposefully designed surveys, where alongside data, the data subjects 'speak for' themselves, and thus the space for interpretation is more transparent if not significantly narrower.

In the context of urban analytics, it is useful to draw on the theoretical perspective on smart urbanism in order to identify the domains in which causal forces may be located. First, we discern the two dimensions involved in the genesis of the data: technology (smartcards) and the subjects (passengers), which in turn are an expression of how the data-capturing technology has been deployed, what passengers do and how technology and subjects interact in everyday life. The technology dimension can be further divided into the cards and the linked network and storage infrastructure that is embedded in a changing transport system. The dimension of data subjects encompasses the time and location of boardings as well as the changing everyday context of senior citizens. Indeed, the reasoning about smart urbanism from a Mobilities perspective suggests that the contextual changes of both transport and senior citizens' lives may be driven by the same technologies and, in turn, the wider socio-economic circumstances governing their deployment; that is smart urbanism as a practice infused with the desire to facilitate and regulate mobility, save cost and increase efficiency within market-oriented styles of governance as part of a wider economic ideology, all of which simultaneously affect the public transport sector, built environments as well as the social, geographical and material living conditions of the elderly.

To summarise, urban analytics, and its salient practical elements of pattern detection, theorisation, contextualisation and interpretation offers potential to generate new and novel strategically relevant hypotheses that would not have been derived without processing of big data; but we note that in none of the studies, big data delivered answers to causal questions. We therefore argue that urban analytics contributes to our understanding of urban systems insofar as it supports the fast generation of novel hypotheses that can be theoretically grounded and contextualised using small, low frequency data.

5. Discussion: urban analytics and planning

What promises does urban analytics hold for urban policy and planning? We have identified the following defining, practical characteristics of urban analytics:

- A focus on real-time, fast dynamics captured in high-frequency data, which contrasts with the slow dynamics of cities' structural changes;
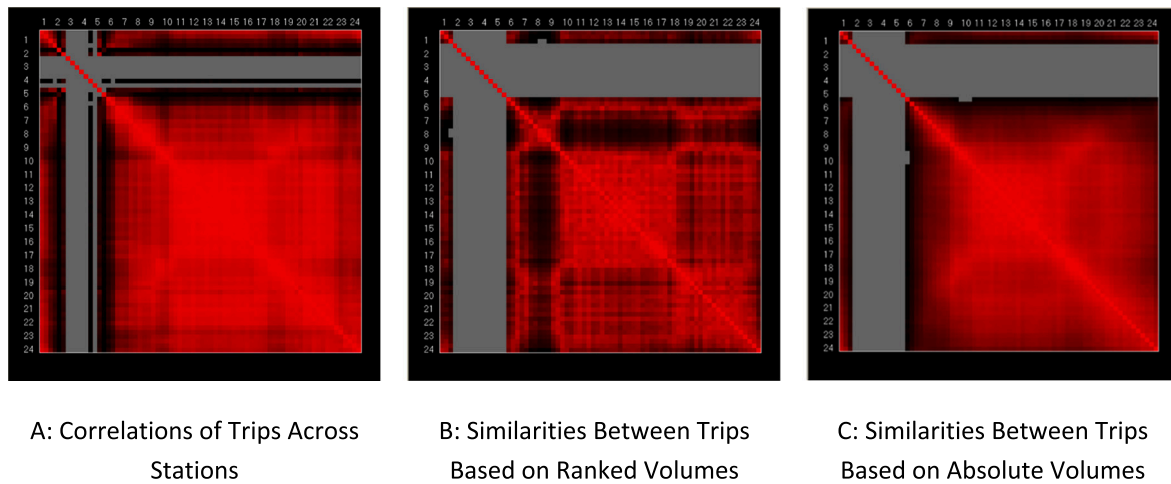
A: Correlations of Trips Across Stations

B: Similarities Between Trips Based on Ranked Volumes

C: Similarities Between Trips Based on Absolute Volumes

Fig. 1. Patterns in the data: correlations and similarities between different time periods.
The grey columns are temporal periods where the statistics are not computed due to closure of the tube during these hours. The range from black to red provides the level of correlation and/or degree of similarity (black low, red high correlation/similarity) where the matrices are obviously symmetric and the main diagonal is computed to give maximum correlation or similarity. The rows and columns are labelled from 00 to 24 h with the size of each row and column being 20-minute intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

- The greater degree and role of subjectivity in interpreting patterns found in big data;
- The political character of the deployment of sensing and computing technologies; and
- The compound nature of captured data in a sense that big data are compound signals of the sensing technology's workings, the activities of data subjects and the mutual impact they have on each other.

With those characteristics in mind, we formulate the following propositions on the value of urban analytics for strategic policy and planning.

1. *Big urban data generate new hypotheses.*
   Big urban data captures activities of both human subjects and sentient environments and are thus like no other type of data source.
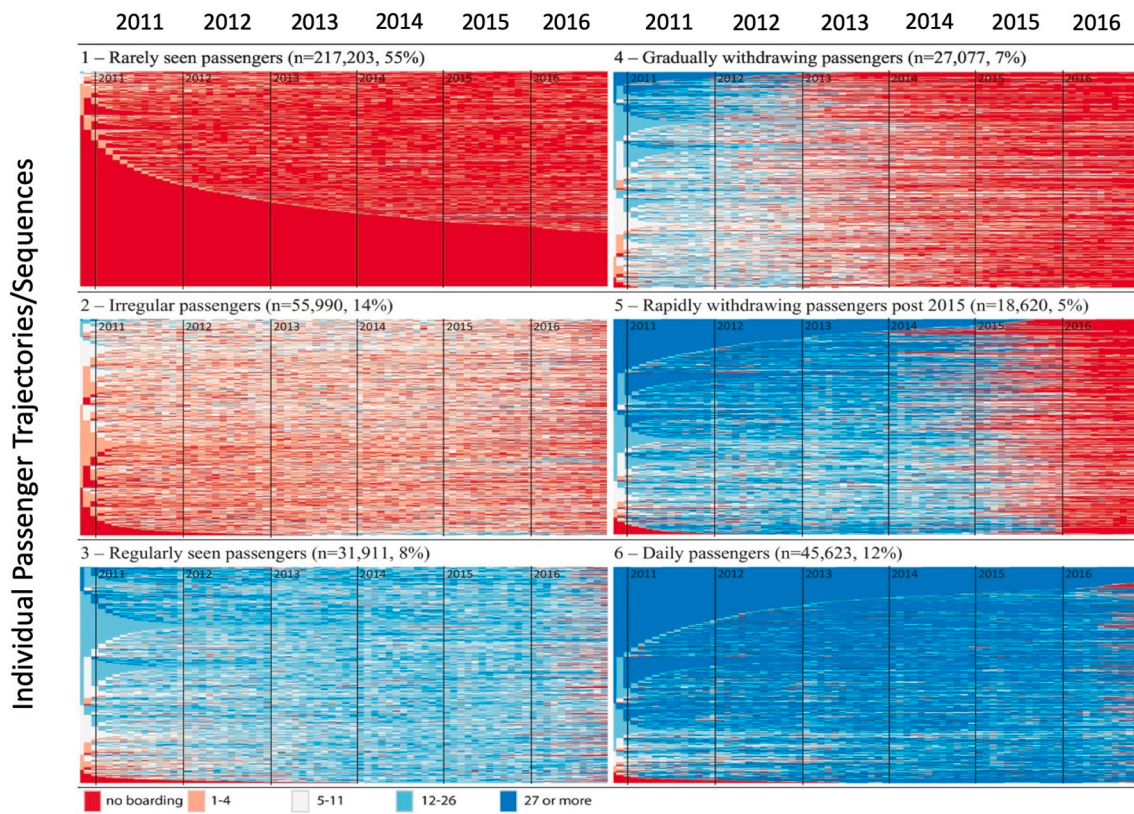

Fig. 2. Six temporal profiles of public transport patronage detected from smartcard data (source: Kandt & Leak, 2019).
In each chart, each line represents an anonymous passenger coloured on a red to blue scale according to the volume of boardings in a four-week time bin. The trajectories are ordered with respect to their relative frequencies of boarding and non-boarding which make the patterns a little clearer. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Hypotheses about the causes of the decline in bus travel by concessionary passholders derived from smart card data spanning six years in the West Midlands, UK.
(Source: Kandt and Leak (2019).)

Hypothesis	Triangulated evidence
1. Online shopping replaces out-of-home trips.	Moderate
2. Senior residents, particularly women, have greater access to the car.	Moderate
3. Greater poverty among pensioners leads to fewer out-of-home activities.	Stronger
4. Premature mortality rates are increasing.	Weak
5. Earlier onset of mobility-limiting disability or ill-health among vulnerable groups.	Weak
6. Cuts in bus routes have reduced accessibility.	Weak
7. Extension of the tram attracts demand away from buses.	Stronger
8. Emergence of new ride-sharing and e-hailing modes attracts demand away from buses.	Stronger

But patterns derived from urban analytics are hardly sufficient to identify a programme of actions for urban policy and planning. As in the case studies, where we can use smartcard data to uncover trends and patterns that would not have been identified otherwise, the patterns as such did not deliver causal insights. Instead, by drawing on geotemporal contextualisation and the theoretical perspective of urban analytics, new hypotheses as to what may cause the patterns could be developed.

We therefore submit that urban analytics rarely if ever provides direct answers to urban policy problems; instead, we would argue that new levels of understanding urban systems can indirectly be attained through using big urban data as exploratory material for the development of new hypotheses. This implies that the major change big data bring to urban research and policy – and perhaps to the social sciences more generally – is a faster pace and new manner of hypothesis generation. Using the new, extended material of big urban data, urban researchers in the big data era may spend more time on formulating and exploring novel causal hypotheses that may be strategically relevant to cities.

2. *Theory becomes more, not less important.*

As the number of datasets and generated patterns increase, theory becomes more crucial than ever in interpreting emergent patterns. Interpretation is a partly cognitive process that necessarily is theory-informed, even if or indeed because it can occur tacitly and subconsciously. A 'lay' daily user of the London Underground is ideally placed to formulate hypotheses about the causes of patterns found in the Oyster card data based on reflexive yet less scientifically structured interpretations of everyday experience. Everyday conscious and unconscious theorisations are necessary human acts as part of practically navigating day-to-day social life (cf Giddens, 1984, 26). We can expect that the same subjective processes are at work when interpreting patterns in big data and thus profoundly influence our ability to make sense of them.

In the two case studies above, the interpretation of patterns and results found in the smart card data required assumptions about the daily mobility practices of urban residents. The tacit 'domain' knowledge could be derived from existing empirical research and practical experience. But in order to transparently identify plausible causal domains, the combined social context of observed data subjects and the data-collecting technology needed to be considered, too. Since the practice-theoretical perspective identifies big urban data as compound signals of human actions, increasingly sentient environments and their open or hidden interactions, we believe that this perspective offers a pertinent framework for structuring interpretation according to a full range of causal domains.

3. *Small data become more, not less important.*

As shown in the case studies, big data do not displace small data; in fact, additional sets of small data are needed to be drawn on to establish the geotemporal characteristics of the patterns (contextualisation), which enabled the generation of hypotheses, and to informally gauge the plausibility of new hypotheses for further

investigation (triangulation). As big urban data enables us to generate more hypotheses faster, our need to test, confirm and reject them expands.

Big data may thus render conventional data sources strategically more important in long-term research, policy and planning, because the certainty arising from controlled, scientifically structured data collection is critical for both contextualisation and triangulation. Although it may well be that traditional research data collection methods, such as sampled surveys, will become less successful (Burrows & Savage, 2014), the use of such sources may still increase as part of big data analytics. Notwithstanding the uncertain future of large-scale surveys, the hypotheses generated from big data can guide new research using established empirical methods, while the contextualised patterns themselves can inform sample designs of quantitative or qualitative inquiries into their causes.

4. *Strategic insights depend on long-term evidence.*

For urban policy that is concerned with causal relations, short-term insights generated from urban analytics need to be framed within long-term trends. In the research case studies, the different patterns associated with a decline in patronage could only be generated from long-term activity records, whereas it is at the resolution of minutes where the patterns could be uniquely characterised by changes in destinations, timing of travel, trip duration to name a few. Rather than understanding causes, which pertain to the 'why', big data then allow an exceptionally refined description of the recorded activities, the 'what'. When we are able to view those fast dynamics over longer time periods, big data undoubtedly make a powerful resource to uncover and characterise deep-seated challenges in contemporary urban systems.

Progress has hitherto been hampered by practical limits with regard to data transfer and storage as well as data sharing restrictions, privacy concerns and commercial or governmental sensitivity (Engin et al., 2020). Although the cost of storage is continuously decreasing, it is our experience that organisations regularly delete historical high frequency data due to limits in storage and processing. Transport for London do not yet systematically use their Oyster card data for any management or policy function and simply archive it as it is collected. While there is extensive work on data visualisation in urban analytics to make big data accessible, the specific challenge in the context of long-term urban policy will be to decide what of the continuously growing data volume should be kept, extracted, summarised or discarded so that they may form assets for strategic decisions in cities. We suggest that, in the context of urban policy, long-term data infrastructures cannot be dissociated from cities' specific political agendas.

5. *Insights from big data rely on contextual analysis.*

In view of the cyclical nature of urban analytics – that is using the same technology that change cities to study them – the motivation of technology deployment and modes of operation need to become integral to the quantitative data analysis. There is little experience with a formal integration of data processing and contextual analysis

of urban sensing technologies, which, instead, tends to occur informally at the stage of interpretation. As we have argued, the theoretical perspective on urban analytics and smart urbanism can help guide the stage of interpretation by starting with the technology-subject nexus to unpack the compound signals of the two dimensions.

In a more formal approach, we may identify and include relevant meta data capturing the operational data. This contextual analysis may include formally codified decisions and events, for example, variables that capture changes to the public transport network, which may have been prompted by observations of demand patterns in smart card data. Such codified 'logs' can then be viewed alongside the thus newly altered patterns in travel demand. Some research collaborations are beginning to explore such approaches (https://mobility.mit.edu/london).

A wider contextual analysis would include the decisions of influential actors, such as private companies, consultancies, academia and city departments who may actively promote the deployment and operation of technologies to ensure continuing return on data infrastructure investments. Such considerations link to the wider critical research agenda of how urban institutions operate, how they deploy the instrumentation of sensors and devices in the urban environment, how they invest in and use the urban informatics platforms the data feed into, and how this affects the kinds of causal reasoning needed in long-term urban policy.

6. *Urban analytics should embrace alternative rationalities.*

Being part of smart urbanism, urban analytics is directly involved in changing cities, most obviously through real-time analysis. Many critics have pointed out that the instrumental rationality that prevails in this form of urbanism is likely to produce the adverse social consequences, including institutionalisation of stigmatisation, social bias and unequal opportunities (Hollands, 2015; Kitchin, 2016b; Rabari & Storper, 2015). We would suggest that the powers of urban analytics – novel characterisation and fast hypothesis generation – can be recast in alternative rationalities to frame novel policy applications.

In urban planning, Alexander (2006) identifies substantive, strategic and communicative rationality in addition to instrumental. Hence, under-explored alternative urban analytics applications include the interim and ex-post evaluations of specific urban policy interventions or the collection of evidence on unintended social consequences of smart urbanism over different time frames. An evaluation of smart urbanism in terms of success on or conflict with other goals and agendas (Martin et al., 2018) would be of both substantive and strategic relevance to cities. Applications based on communicative rationality could focus on the degree to which analytics-powered characterisation and visualisation can engage a wider range of civil society members in city-making. More conceptual and empirical work on urban analytics applications that embrace non-instrumental rationalities are needed; and we would argue that such work alongside reflexive, theoretical framing grounded in contemporary, accelerated forms of Mobilities and smart urbanism as a practice are necessary steps to enable urban analytics to make positive contributions to urban policy and planning in terms of a social, egalitarian vision.

6. Concluding remarks

Although it remains uncertain how the practice of urban analytics will influence the way we might plan cities, social science critiques demonstrate that, if data-driven urban policy is enacted through instrumental rationality, automated 'software-sorting' will become fundamental to organising cities (Kitchin & Dodge, 2011; Wood & Graham, 2006). This reshaping of urban institutions is likely to reproduce existing power relations and amplify the influence of technical experts, as reliance on increasingly complex, physico-digital systems increases

(Kitchin, 2019).

Against this social and political backdrop, we identified important epistemological and practical challenges that arise when high-frequency data are analysed for purposes of strategic urban policy. We found that big data and urban analytics advance our understanding of urban systems insofar as they generate new and novel hypothesis perhaps at a faster pace than before. Concurring with Singleton and Arribas-Bel (2019), we would argue that the plausibility, relevance and thus strategic value of hypotheses materialise through transparent theoretical grounding and geotemporal contextualisation. In this process, the importance of long-established, low-frequency knowledge assets will increase not despite but because of the rise of big data.

It is clear that there is further need for conceptual and empirical work on urban analytics that addresses the epistemological, practical and normative challenges that arise in the context of urban policy. In discussing our six propositions on urban analytics, we identified a number of research needs that would usefully contribute to the discussion.

First, more attention should be paid to the cognitive processes that are involved in interpreting patterns found in big data within specific institutional settings. As we have shown, the room for and discretion in interpretation is larger than in survey data, where data subjects actively and discursively provide information as part of a scientifically designed inquiry. Thus, theoretical reasoning and steps to contextualise play a much greater yet more elusive role in the practice of big data analytics.

Second, in view of the regulatory nature of sensing and computing technologies, more research is needed on the ways in which the technologies that enable urban analytics are invested in and deployed in particular social, geographical and temporal contexts. But the contextual sensitivity of big data also implies that urban analytics should expand the focus from what we can learn about human 'data' subjects to human-technology interactions, their institutionalisation and the character of new socio-technical assemblages. In other words, urban analytics should abandon its often behaviourist orientation and focus on the compound signals of human actions, institutional, technical and regulatory characteristics recorded in increasingly sentient and responsive environments. Such reorientations may be accompanied by conceptual work that identifies ways of urban analytics to adopt non-instrumental types of rationality to inform new and novel policy actions addressing long-term social challenges facing cities.

Finally, research should advance a fuller framing of the fast dynamics recorded in real-time within long-term trends in cities. Big urban data will soon be available over long-term periods eliminating issues of aggregation and temporal scale. We will then enter an era where short and long term begin to merge and this may have major implications for the theories and methods we adopt in our understanding of cities. But inasmuch as opportunities for data mining and hypothesis-generation over larger time frames increase, so too does the need for contextualisation and triangulation. Integrating all forms of data – fast and slow, big and small, quantitative and qualitative – will be an essential step in promising applications of big data in urban policy and planning.

Declaration of competing interest

The authors have no conflicts of interests.

Acknowledgements

References

Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2017). A model for the analysis of data-driven innovation and value generation in smart cities' ecosystems. *Cities, 64*, 47–53. https://doi.org/10.1016/j.cities.2017.01.011.

Albino, V., Berardi, U., & Dangelico, R. M. (2015). Smart cities: Definitions, dimensions, performance and initiatives. *Journal of Urban Technology, 22*(1), 3–21.

Alexander, E. R. (2006). *Evaluation in Planning: Evolution and Prospects.* Aldershot UK: Ashgate Publishing Ltd.

Anda, C., Erath, A., & Fourie, P. J. (2017). Transport modelling in the age of big data. *International Journal of Urban Sciences, 21*, 19–42. https://doi.org/10.1080/12265934.2017.1281150.

Babar, M., & Arif, F. (2017). Smart urban planning using Big Data analytics to contend with the interoperability in Internet of Things. *Future Generation Computer Systems, 77*, 65–76. https://doi.org/10.1016/j.future.2017.07.029.

Barns, S., Cosgrave, E., Acuto, M., & Mcneill, D. (2017). Digital infrastructures and urban governance. *Urban Policy and Research, 35*(1), 20–31. https://doi.org/10.1080/08111146.2016.1235032.

Batty, M. (1997). Virtual geography. *Futures, 29*(4–5), 337–352.

Batty, M. (2018). *Inventing future cities.* Cambridge, MA: The MIT Press.

Batty, M. (2019). Urban analytics defined. *Environment and Planning B: Urban Analytics and City Science, 46*(3), 403–405.

Batty, M. (2020). Defining smart cities: High and low frequency cities, big data and urban theory. In K. Willis, & A. Aurigi (Eds.). *The Routledge companion to smart cities.* London: Routledge.

Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., & Portugali, Y. (2012). Smart cities of the future. *European Physical Journal: Special Topics, 214*(1), 481–518.

Briand, A. S., Côme, E., Trépanier, M., & Oukhellou, L. (2017). Analyzing year-to-year changes in public transport passenger behaviour using smart card data. *Transportation Research Part C: Emerging Technologies, 79*, 274–289. https://doi.org/10.1016/j.trc.2017.03.021.

Boeing, G., Batty, M., Jiang, S., and Schweitzer, L. (2020) Urban analytics: History, trajectory, and critique, in S. Rey and R. Franklin (editors) Handbook of spatial analysis in the social sciences, Edward Elgar, Cheltenham, Glos, UK.

Bok, R., & Coe, N. M. (2017). Geographies of policy knowledge: The state and corporate dimensions of contemporary policy mobilities. *Cities, 63*, 51–57.

Bourdieu, P. (1977). *Outline of a Theory of Practice (Vol. 16, Cambridge Studies in Social and Cultural Anthropology).* New York: Cambridge University Press.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information Communication and Society, 15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878.

BS ISO 37106:2018 (2018). *Sustainable cities and communities. Guidance on establishing smart city operating models for sustainable communities.* https://shop.bsigroup.com/ProductDetail/?pid=000000000030348126.

Bunders, D. J., & Varró, K. (2019). Problematizing data-driven urban practices: Insights from five Dutch 'smart cities'. *Cities, 93*, 145–152.

Burrows, R. J., & Savage, M. (2014). After the crisis? Big Data and the methodological challenges of empirical sociology. *Big Data and Society, 1*(1), 1–6.

Campbell, T. (2012). *Beyond smart cities: How cities network, learn and innovate.* London: Earthscan.

Castells, M. (1996). *The rise of the network society.* London: Wiley-Blackwell.

Chu, K. K. A. (2015). Two-year worth of smart card transaction data - Extracting longitudinal observations for the understanding of travel behaviour. *Vol. 11.* https://doi.org/10.1016/j.trpro.2015.12.031.

Delmelle, E. C. (2019). Toward a more socially impactful geographical analysis. *Geographical Analysis, 1–9.*

Dijkstra, E. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik, 23*(3), 269–271.

Engin, Z., et al. (2020). Data-driven urban management: Mapping the landscape. *Journal of Urban Management, 9*(2), 140–150.

Gaffney, C., & Robertson, C. (2016). Smarter than smart: Rio de Janeiro's flawed emergence as a smart city. *Journal of Urban Technology, 25*(3), 47–64. https://doi.org/10.1080/10630732.2015.1102423.

Gauthier, JA., Bühlmann, F., & Blanchard, P. (2014). Introduction: Sequence Analysis in 2014. In Bühlmann F. Blanchard P.Gauthier JA (Ed.). *Advances in Sequence Analysis: Theory, Method, Applications. Life Course Research and Social Policies, vol 2*Springer, Chamhttps://doi.org/10.1007/978-3-319-04969-4_1.

Gesis (n.d.). Computational social science. Leibniz Institute for the Social Sciences, www.gesis.org/en/institute/departments/computational-social-science .

Giddens, A. (1984). The constitution of society. *Cambridge UK.* Malden US: Polity Press.

Goulet Langlois, G., Koutsopoulos, H. N., & Zhao, J. (2016). Inferring patterns in the multi-week activity sequences of public transport users. *Transportation Research Part C: Emerging Technologies, 64*, 1–16. https://doi.org/10.1016/j.trc.2015.12.012.

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

Graham, S., & Marvin, S. (1996). *Telecommunications and the city: Electronic spaces, urban places.* London: Routledge.

Graham, S., & Marvin, S. (2001). *Splintering urbanism: Networked infrastructures, technological mobilities and the urban condition.* London: Routledge.

Greenfield, A. (2013). *Against the smart city.* New York: Do Projects.

Greenfield, A. (2017). *Radical technologies: The design of everyday life.* London: Verso.

Harrison, C., Eckman, B., Hamilton, R., Hartswick, P., Kalagnanam, J., Paraszczak, J., & Williams, P. (2010). Foundations for smarter cities. *IBM Journal of Research and Development, 54*(4), 1–16. https://doi.org/10.1147/JRD.2010.2048257.

Hollands, R. G. (2008). Will the real smart city please stand up? Intelligent, progressive or entrepreneurial? *City, 12*(3), 303–320. https://doi.org/10.1080/13604810802479126.

Hollands, R. G. (2015). Critical interventions into the corporate smart city. *Cambridge Journal of Regions, Economy and Society, 8*(1), 61–77. https://doi.org/10.1093/cjres/rsu011.

Huang, J., Levinson, D., Zhou, J., Wang, Z., & Wang, J. (2018). Tracking job and housing dynamics with smartcard data. *Proceedings of the National Academy of Sciences, 115*(50), 12710–12715. https://doi.org/10.1073/pnas.1815928115.

Kandt, J., & Leak, A. (2019). Examining inclusive mobility through smartcard data: What shall we make of senior citizens' declining bus patronage in the West Midlands? *Journal of Transport Geography, 79*, 102474 (online pre-print).

Kang, W., Oshan, T., & Wolf, L. J. (2019). *Defining urban data science: A roundtable discussion.* Environment and Planning B: Urban Analytics and City Science46 (in press).

Kitchin, R. (2014a). Big Data, new epistemologies and paradigm shifts. *Big Data & Society, 1*(1), 1–12. https://doi.org/10.1177/2053951714528481.

Kitchin, R. (2014b). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1–14. https://doi.org/10.1007/s10708-013-9516-8.

Kitchin, R. (2016a). Reframing, reimagining and remaking smart cities. *SocArXiv papers.* https://doi.org/10.31235/osf.io/cyjhg.

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

Kitchin, R. (2019). The timescape of smart cities. *Annals of the American Association of Geographers, 109*(3), 775–790. https://doi.org/10.1080/24694452.2018.1497475.

Kitchin, R., & Dodge, M. (2011). *Code/space : Software and everyday life.* London: MIT Press: Cambridge, Mass.

Kitchin, R., Coletta, C., Evans, L., Heaphy, L., & Macdonncha, D. (2017). Smart cities, epistemic communities, advocacy coalitions and the "last mile" problem. *Information Technology, 59*(6), 275–284.

Kitchin, R., Lauriault, T. P., & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science, 2*(1), 6–28. https://doi.org/10.1080/21681376.2014.983149.

Komninos, N. (2011). Intelligent cities: Variable geometries of spatial intelligence. *Intelligent Buildings International, 3*(3), 172–188.

Komninos, N., Pallot, M., & Schaffers, H. (2013). Special issue on smart cities and the future internet in Europe. *Journal of the Knowledge Economy, 4*(2), 119–134.

Kong, L., & Woods, O. (2018). The ideological alignment of smart urbanism in Singapore: Critical reflections on a political paradox. *Urban Studies, 55*, 679–701. https://doi.org/10.1177/0042098017746528.

Kourtit, K., Nijkamp, P., & Steenbruggen, J. (2017). The significance of digital data systems for smart city policy. *Socio-Economic Planning Sciences, 58*, 13–21. https://doi.org/10.1016/j.seps.2016.10.001.

Larsen, J., Urry, J., & Axhausen, K. W. (2006). *Mobilities, networks.* Ashgate, Aldershot, UK: Geographies.

Lash, S. (2002). *Critique of information.* London: Sage Publications.

Liu, X., Zhou, Y., & Rau, A. (2019). Smart card data-centric replication of the multi-modal public transport system in Singapore. *Journal of Transport Geography, 76*(December 2017), 254–264. https://doi.org/10.1016/j.jtrangeo.2018.02.004.

Luque-Ayala, A., & Marvin, S. (2015). Developing a critical understanding of smart urbanism? *Urban Studies, 52*(12), 2105–2116. https://doi.org/10.1177/0042098015577319.

Martin, C. J., Evans, J., & Karvonen, A. (2018). Smart and sustainable? Five tensions in the visions and practices of the smart-sustainable city in Europe and North America. *Technological Forecasting and Social Change, 133*(December 2017), 269–278. https://doi.org/10.1016/j.techfore.2018.01.005.

Merics (2017). China's social credit system. A big-data enabled approach to market regulation with broad implications for doing business in China. Retrieved from Mercator Institute for China Studieshttps://www.merics.org/sites/default/files/2018-03/171212_China_Monitor_43_Social_Credit_System_Implementation_1.pdf.

Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current trends in smart city initiatives: Some stylised facts. *Cities, 38*, 25–36. https://doi.org/10.1016/j.cities.2013.12.010.

Peck, J., & Theodore, N. (2010). Mobilizing policy: Models, methods, and mutations. *Geoforum, 41*(2), 169–174. https://doi.org/10.1016/j.geoforum.2010.01.002.

Pelletier, M., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C, 19*(4), 557–568. https://doi.org/10.1016/j.trc.2010.12.003.

Prince, R. (2015). The spaces in between: Mobile policy and the topographies of the technocracy. *Environment and Planning D: Society and Space, 34*(3), 420–437.

Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Mary Ann Liebert Inc. 1*(1), 51–59.

Rabari, C., & Storper, M. (2015). The digital skin of cities: Urban theory and research in the age of the sensored and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions, Economy and Society, 8*(1), 27–42.

Rathore, M. M., Ahmad, A., Paul, A., & Rho, S. (2016). Urban planning and building smart cities based on the Internet of Things using Big Data analytics. *Computer Networks, 101*, 63–80. https://doi.org/10.1016/j.comnet.2015.12.023.

Reades, J., Zhong, C., Manley, E., Milton, R., & Batty, M. (2016). Finding pearls in London's oysters. *Built Environment, 42*(3), 365–381.

de Santana, E. S., Nunes, E. O., Costa Passos, D., & Santos, L. B. (2019). SMM: A maturity model of smart cities based on sustainability indicators of the ISO 37122. *International Journal of Advanced Engineering Research and Science, 6*(2), 13–20.

Sheller, M. (2017). From spatial turn to mobilities turn. *Current Sociology, 65*(4), 623–639. https://doi.org/10.1177/0011392117697463.

Sheller, M., & Urry, J. (2006). The new Mobilities paradigm. *Environment and Planning A,*

38(2), 207–226.

Singleton, A. D., & Arribas-Bel, D. (2019). Geographic data science. *Geographical Analysis,* 1–15.

Söderström, O., Paasche, T., & Klauser, F. (2014). Smart cities as corporate storytelling. *City, 18*(3), 307–320. https://doi.org/10.1080/13604813.2014.906716.

Storme, T., Faulconbridge, J., Beaverstock, J., Derudder, B., & Witlox, F. (2017). Mobility and professional networks in academia: An exploration of the obligations of presence. *Mobilities, 12*(3), 405–424.

Taylor, L. (2015). *No place to hide?* The ethics and analytics of tracking mobility using mobile phone data. Environment and Planning D: Society and Space3. https://doi.org/10.1177/0263775815608851.

Taylor, Linnet (2016). No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environment and Planning D: Society and Space*. https://doi.org/10.1177/0263775815608851.

TfL (2018). Delivering the next generation of urban traffic management. retrieved from Transport for Londonhttps://tfl.gov.uk/info-for/media/press-releases/2018/june/delivering-the-next-generation-of-urban-traffic-management.

Thrift, N., & French, S. (2002). The automatic production of space. *Transactions of the Institute of British Geographers, 27*(3), 309–335.

Townsend, A. M. (2013). *Smart cities: Big data, civic hackers, and the quest for a new utopia.* Norton, New York: W. W.

Urry, J. (2000). *Sociology beyond societies: Mobilities for the Twenty-first Century.* London: Routledge.

Urry, J. (2007). *Mobilities.* Cambridge, UK: Polity.

Urry, J. (2012). Social networks, mobile lives and social inequalities. *Journal of Transport Geography, 21*, 24–30. https://doi.org/10.1016/j.jtrangeo.2011.10.003.

Vanolo, A. (2014). Smartmentality: The smart city as disciplinary strategy. *Urban Studies, 51*(5), 883–898. https://doi.org/10.1177/0042098013494427.

WEF (2018). The Known Traveller. Unlocking the potential of digital identity for secure and seamless travel. Retrieved from World Economic Forumhttps://www.weforum.org/reports/the-known-traveller-unlocking-the-potential-of-digital-identity-for-secure-and-seamless-travel.

Wiig, A. (2015). IBM's smart city as techno-utopian policy mobility. *City, 19*(2–3), 258–273. https://doi.org/10.1080/13604813.2015.1016275.

Wood, D. M., & Graham, S. (2006). Permeable boundaries in the software-sorted society: Surveillance and differentiations of mobility. In M. Sheller, & J. Urry (Eds.). *Mobile technologies of the city*. Abingdon, UK: Routledge.

Wyly, E. (2014). Automated (post)positivism. *Urban Geography, 35*(5), 669–690. https://doi.org/10.1080/02723638.2014.923143.

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science, 28*(11), 2178–2199. https://doi.org/10.1080/13658816.2014.914521.

Zhong, C., Batty, M., Manley, E., Wang, J., Wang, Z., Chen, F., & Schmitt, G. (2016). Variability in regularity: A comparative study of urban mobility patterns in London, Singapore and Beijing using smart-card data. *PLoS One, 11*(2), Article e0149222. https://doi.org/10.1371/journal.pone.0149222.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for the future at the new frontier of power.* London: Profile Books.

Zygiaris, S. (2013). Smart city reference model: Assisting planners to conceptualize the building of smart city innovation ecosystems. *Journal of the Knowledge Economy, 4*(2), 217–231.