palgrave communications
HUMANITIES | SOCIAL SCIENCES | BUSINESS

ARTICLE

https://doi.org/10.1057/s41599-019-0264-3 OPEN

The data politics of the urban age

Fábio Duarte[1,2] & Ricardo Álvarez[1]

ABSTRACT The deployment of myriad digital sensors in our physical environments is generating huge amounts of data about the natural and built environments and about ourselves, social relations, and interactions in space. These unprecedented quantities of data combine with high-performance computers to produce a series of increasingly powerful tools ranging from mathematical modeling on a massive scale to various types of artificial intelligence. Within this context, urban planning and design driven by data and predictive tools have been gaining traction. This scientific approach to urban problems echoes the nineteenth-century birth of modern urbanism, when rapid industrialization and new scientific methods were advocated against a traditional beaux-arts approach to city planning; and the twentieth century proved that such scientific methods were politically charged. Arguing that we are facing a similar breakthrough in urban studies and planning, in this paper we discuss how data-driven approaches can foster urban studies, but must be balanced with a critical view to the inherent social values of cities.

[1] Massachusetts Institute of Technology, Cambridge, MA, United States. [2] Pontificia Universidade Catolica do Parana, Curitiba, Brazil. Correspondence and requests for materials should be addressed to F.D. (email: fduarte@mit.edu)

Introduction

If a term can define an era, ours is the urban age. Cities today are inhabited by more than 50% of the world's population, who are responsible for 75% of the planet's energy consumption and for 80% of CO_2-related emissions. The global urban population will continue to rise dramatically in the near future, with some projections targeting a further increase of an additional 2.6 billion people within 30 years' time (United Nations, 2015). This means that many of the environmental, economic and social problems of our era are and will continue to be primarily urban challenges; after all, urban dynamics are not independent from the natural inflows and outflows of matter, energy, and information. Development of tools to map and measure these flows is therefore critical to create a comprehensive understanding of the urban phenomenon at different scales, as well as to inform policy and design interventions more efficiently.

Fortunately, there is an unprecedented abundance of data currently generated in cities. The widespread use of informational infrastructures and personal devices people use to communicate with each other and access third-party web-based services leaves behind trillions of digital traces of activities and interactions, often with additional metadata such as timestamps and geolocations which are very useful for further analysis. Wearable devices with onboard biometric sensors measure personal health conditions and physical activities, while sensors measure changing features of the built and the natural environment, all constantly amassing data. The modern digital layers that feed our daily lives generate vast troves of data, which collectively represent the signature of human relations with our environment. This represents an opportunity to understand urban phenomena from new perspectives, since most of this data was not available before; one of the questions we discuss in this paper is whether the ongoing research has been addressing the questions and issues that are relevant for policy makers and urban planners.

The combination of huge volumes of data and machines with increasing computational power has fed the growth of both, while pushing for new analytical tools in the process. Warren Weaver's prediction that computational power and multidisciplinary approaches would need to be harnessed to answer questions of organized complexity has come to pass (Shannon and Weaver, 1969). Gradually, the wide availability of data and cheaper FLOPS (floating point operations per second, a measure of processing power) have shifted our methods of understanding complexity from mere statistical analysis based on a limited number of samples aiming to validate larger systems to in-depth analysis that simultaneously measures the entire universe of available data; the new goal is not to infer validity, but rather to state a measurable reality, creating a deep impact in the fundamental processes of scientific discovery. Meanwhile, as the complexity of our systems and their related data exceed our biological capacity of understanding and abstraction, we have seen the recent emergence of machine learning methods such as deep learning (essentially Bayesian machines with a recursive learning cycle) which are able to leverage modern computational capabilities to establish interrelations within extremely complex systems hidden in vast volumes of data points and interconnect them into information systems of aggregated complexity that function on nanosecond scales.

The vast availability of urban data makes cities a clear testing ground for artificial intelligence technologies. Machine learning methods ranging from convolutional neural networks to deep neural networks, deep cascades, and more recently generative adversarial networks are increasingly used in urban planning and urban tech (an area that includes design, research, and development of technology applied to solve urban problems) to recognize patterns hidden in large volumes of multidimensional data. Such data may include audio and video collected by individuals or governments and companies, mobile phone communication data, lidar sensor data from autonomous vehicles, environmental data collected by organizations or individuals, and human-related biological data we disperse in our daily activities when we use toilets or pass through biometric sensors at airports, office buildings, or even some refugee camps, to give only a few examples. Taking advantage of these troves of data, data scientists have been using artificial intelligence technologies to reveal aspects of city life which have never been captured before, as well as to lay the foundation of technologies that could change the way we understand, optimize, design, and ultimately live in cities. Now it is time to translate such findings into innovative policy and planning approaches. On the one hand, it will require data scientists to understand that urban issues not always translate into datasets, but on the other hand, it will demand an effort from planners and policy makers to accept and understand these new methodologies—and some unexpected findings.

This is not a new dynamic, in the historical sense. Over one hundred and fifty years ago, rapid technological transformations triggered profound changes that drove more people from rural areas to booming urban economies. Architects and planners faced a somewhat similar problem: how to make sense of contemporary urbanization and plan and design cities when the established methods do not respond accordingly? Immersed in an era when *beaux-arts* principles dictated what a good city should be, a few urban planners appropriated emerging scientific tools and methods and advocated for data-driven approaches to urban challenges. Ildefons Cerdà expressed this scientific approach to urban planning in his writings, and more importantly, in his master plan for Barcelona. His plan was based on extensive measurements of the physical and social aspects of the city, and he argued that outcomes should be evaluated in terms of better traffic, healthy households, air flows in urban canyons, and overall hygienist principles. This was the emergence of modernist urbanism, in which accurate measurements, scientific analysis, and efficiency became paramount and drove design.

As in any scientific field, it is only possible to understand what can be detected and measured, the measurements always depend on the tools available, and the findings depend on available analytical methods. These constraints eventually narrow what is accepted as belonging to the field, and any result that does not follow these procedures is usually not taken seriously. Each of these assumptions (what can be measured, what tools are appropriate, what methods are valid) embodies broader political decisions that extend far beyond the laboratories and the specific field. This is even more applicable when it comes to cities: living environments made of multiple natural, human, and man made physical, symbolic, and institutional components which are constantly interrelating.

In the 1960s, a movement countering the "scientific" approach to urban planning and design arose, bringing cultural, political, and social aspects back to urbanism. Advocacy groups and participatory methods were introduced as key components of planning, along with topics such as environmental justice, cultural heritage, and social values. Yet this rebalancing took place at a time without any major technological breakthroughs comparable to those seen today, when digital technologies pervade virtually all aspects of our lives and new tools and methodologies allow us to detect, measure, and analyze phenomena that were impossible to fathom twenty years ago. We may again be on the verge of a paradigm shift in terms of how we understand urban phenomena and propose solutions to complex urban challenges. In this paper we argue for a critical perspective in relation to these data-driven approaches, highlighting the politics inherent in any technological and scientific method, particularly when focus of research subject is our complex, living, ever-evolving cities. Since this is not the first wave of data-driven or technology-driven or science-driven approach to urban challenges, it is important to analyze this with a balance between openness to what is actually new and transformative, but also some skepticism toward findings that might be clever but do not address the real and pressing urban issues.

First, we argue that such technological changes occur within a phenomenon of urbanization which is happening on a global scale. Second, we discuss data-driven approaches which rely on the abundance of data collected by pervasive sensors. Third, we suggest that both a critical approach to big data, as well as policy and design innovations might take advantage of the scaling laws that characterize urban phenomena. In these three sections, we argue that however powerful, these approaches might inadvertently favor data-rich cities, excluding large swaths of the urban world that are still data deserts. Finally, within this critical perspective, we explore how these innovations can inform policy and design innovations.

Urbanization on multiple scales

Scholars from social sciences and geography (Lefebvre, 1970; Santos, 1985) have argued for decades that the definition of *urban* should not be confused with the physical form urbanization usually takes (a densely populated built area). Rather, urban should be understood as a phenomenon involving economic and social relations, scientific and technological innovations, and cultural and political shifts that are triggered and promoted by the concentration of social interactions in a relatively constrained space. Regardless of the size of any specific city, the urban phenomenon eventually reaches a global scale.

Within this context, cities (and particularly major cities, with the power agglomeration brings to a number of activities by increasing and strengthening the interactions between them) are the prevalent paradigm through which the modern world can be understood. In this context, urban is conceived as a "concrete abstraction," where socio-spatial relations are simultaneously territorialized (concrete) and generalized (abstraction), extending across multiple scales, from localities to the world (Brenner, 2013). But as Neil Brenner (2018) insists, planetary urbanization does not imply that it flattens or supersedes local differences; urbanization is not a city-contained phenomenon, and cannot be fully understood if the city, frequently seen as a bounded space, is considered the epistemological starting point.

Therefore, analysis of the environmental and social dynamics that occur at the city level needs to be balanced with the understanding of the global reach of the urban phenomenon. Tackling urban problems "demands robust, sophisticated and truly global urban research" (Acuto et al., 2018), research which must be evidence-based, scientifically consistent, and utilize clear and replicable methodologies. This approach would arguably stimulate data-driven decision-making processes, producing impacts than can be measured in the long term. Data-driven research would consequently lead to evidence-based design solutions and make urban governance more accountable, at the local, as well as at the global (and comparative) level.

One of the challenges in strengthening urban governance driven by multidisciplinary research-based findings lies in the epistemological gap that still exists between science and polity. As we have been pointing out here, this will require policy makers to be prepared and open to methods that are new to the field (too any field, actually), but also data scientists to accept that many urban issues do not translate to datasets.

If we consider urbanization on the global scale, even data-driven city governance faces the "metrocentric" challenge (Bunnell and Maringanti 2010; Acuto, 2018), in which abundant data are collected and analyzed in global metropolises from São Paulo to London to Singapore, while the majority of medium-sized and small cities (particularly in the Global South) do not collect any data, and do not have the skill sets to deal with data-driven urban research. This is not a trivial situation, since far more cities around the world share the dynamics and realities of medium-sized and small cities like Medellin, Chiang Mai, and Benguerir than large metropolises like Paris, London, or New York. The majority of the world's population and the main drivers of population and urban growth currently reside (and will continue to reside) in medium-sized and small cities, making the dearth of data and related urban knowledge a painful gap.

Indeed, in recent years cities have gained increasing relevance as the drivers of social, political, economic, and environmental decisions. Because the city is the geographic scale most closely related to people's daily lives, making the city the center of global issues is a strong political statement in itself. We must be careful, though. While advocacy for city-centric governance that bypasses nation-states has captured the attention of international agencies like the United Nations, as well as large private conglomerates, an interest in placing cities at the forefront of policy-making could also be a socially acceptable way of pushing particular economic agendas without disrupting broader political arrangements.

With these precautions in mind, cities undeniably attract people and generate cultural, scientific, and technological innovations. Since many societal struggles and innovations take place in urban environments, cities are and will continue to be drivers of change on a global scale.

Cities as data generators

As the centers of gravity for human activity, cities are the key data generators of our times. The pervasiveness of sensors collecting data from individuals and social groups, natural and built environments, and the relationships between these elements are creating voluminous datasets; when these banks of data are coupled with artificial intelligence techniques they give rise to new possibilities for thinking about urban phenomena and proposing design and policy solutions. The challenge we discuss here is the necessary balance between data-driven approaches with socially critical views.

In ten years, 150 billion networked measuring sensors will be deployed in the world (Helbing, 2019). The sensor market has grown over 10% per year in recent decades, boosted by increasing interest in the Internet of Things (IoT), a platform for devices to exchange data, make decisions, and act without direct human interference. Although it is not the only one, IoT is a key advancement by which the physical and the digital layers of the city merge.

Sensors and IoT devices as data generators and transactors are only two pieces of this puzzle. In general terms, there are three basic types of data gathering: purposely sensed, user generated, and opportunistic. Purposely sensed data comes from sensors designed and deployed for specific uses, which can be used in related situations. An example is thermal cameras deployed on vehicles to detect building heat loss, temperature, and air quality in cities (Anjomshoaa et al., 2018). User generated describes data which is purposely generated and made available by humans through direct interaction with their bodies and minds, such as portable electroencephalography readers and self-tracking devices (like walking and biking apps) (Vanky et al., 2017), as well as data shared by users through digital platforms such as social media, online activities, and specific utilization of mobile applications.

This data can be aggregated to inform, as well as engage citizens in decision-making processes. Finally, opportunistic data gathering takes data that was originally generated and collected for a specific purpose and uses it for other purposes outside its original design. For example, cell phone data used for communication between individuals can be aggregated and used to map mobility patterns on a large scale with fine temporal and spatial granularity (Calabrese et al., 2011). The key point here is to utilize available data for purposes beyond its original design and intent; this can be applied to data from sensors and digital devices, as well as data generated by humans.

But what happens when and where data is not available? Theories and solutions are based on data. If data doesn't exist for a particular area or population, these theories and solutions are inherently biased. To provide an example, Kessler et al., (2016) have recently shown that breakthrough medical treatments have a heavy racial bias, with European-based data predominating within clinical databases. This bias in the dataset determines which genetic diseases prevalent within this population will get more research funding and attention from the pharmaceutical industry, and eventually be treated. When it comes to urban challenges, there is an evident technological bias: current methods depend on the abundance of data, which is only generated, gathered, or transmitted with the help of digital technologies.

What can be done is to identify critical interdependencies among separate sub-systems (Ang et al., 2017). This follows a critical characteristic of data: its potential value invariably grows when combined with other data. Indeed, as discussed before, innovative solutions for urban problems must consider that the urban phenomenon, as well as data generation in urban environments are reaching global scales, taking advantage of the ease with which digital systems grow. Therefore, in order to employ data science to inform public policies and urban design, we need methodological tools to understand the global scale of urbanization and their transferability within local contexts and dimensions. One example is recent interest in the properties of scaling laws as epistemological tools for understanding global urbanization, which we discuss below.

Big data and the scaling laws of cities

The use of big data in urban studies can take advantage of the natural scalability laws that characterize cities within similar morphological, socio-economic, and political contexts. Although each city has its peculiarities, comparative studies involving hundreds of cities have demonstrated that infrastructural, social, and economic features respect scaling laws, with interaction-related variables following super-linear scalings, while infrastructure variables follow sub-linear scalings and household-related variables follow nearly linear scalings (Li et al., 2017). When cities grow (in terms of population, for instance) more social interactions occur, and super-linear scalability is perceived in everything from wages to diseases, from the number of patents filed to criminality.

The larger the city, the more social interactions occur over space and time, and the more the average citizen owns, produces, and consumes. Scaling exponents are similar across several European countries, which can easily be seen using double-logarithmic representations (Kühnert et al., 2006). As Geoffrey West (2017, p. 278) states,"within their own urban systems [cities] are approximately scaled versions of one another." The growing literature on the scaling properties of urban systems has not yet led to a theory which can guide how cities should be planned, but particular areas (such as land-use-transportation models) have been incorporating these ideas in assessing the efficacy and general consequences of such plans (Batty, 2008).

Again, a critical perspective is necessary. These methods have yielded promising results based on large amounts of data, but such datasets can only be obtained where this data is collected. Consequently, a key challenge is how to address geographical imbalances of data availability worldwide. In a data-driven society, the solutions to contemporary urban problems are most likely to emerge first in locations with data-rich ecosystems. Meanwhile, data deserts (composed of large swaths of the urban world formed by small cities in developing nations where data is seldom gathered and made available) will be deprived of such solutions for longer periods. Beyond the technological challenges, there are also political obstacles which must be confronted in order to reframe the scientific agenda on urban issues. As Acuto et al. (2018) argue, as in other fields, small pockets of data-rich areas are more likely to receive research funding and address topics of interest to policy and market drivers—and these themes include smart cities.

In order to foster data-driven participatory processes and urban solutions, several groups have advocated for open data policies as part of smart cities initiatives (Ching and Ferreira 2015; Attard et al., 2015). But a question that underlies open data initiatives is seldom addressed: who benefits from open data? While open data is usually promoted as a way to provide everyone equal opportunities to access data, the reality is that the skills and resources required to effectively generate new knowledge that can be transformed into innovations create a barrier across different social groups, some of which may be disadvantaged in terms of reaping benefits. The scientific community in the Global South may be particularly vulnerable to this phenomena.

Making sense of data: artificial intelligence and cities

Leveraging the data-driven capabilities required to understand and operationalize systems of even greater complexity and scale increasingly involves a variety of machine learning (ML) and artificial intelligence (AI) techniques. Artificial intelligence, succinctly defined by Merriam-Webster[1] dictionary [tm6] as "the capability of a machine to imitate intelligent human behavior," is a sub-field of computer science, which for decades has attempted to reach this goal through a variety of approaches ranging from decision trees and expert systems to neural networks and probabilistic machines. After the first "AI" winter of the 1970s and early 1980s, when research funding dried up and expert systems were becoming unmanageable and expensive, machine learning effectively took off as the paradigm for getting a computer to come up with solutions to complex problems based on identifiable patterns in multidimensional data.

Today's power of computers to make high-level abstractions by learning through available data is fundamentally driven by deep learning methods (Lecun et al., 2015) that use a back-propagation function to refine results at each data iteration. These methods involve interconnection of multiple (deep) layers of artificial neural networks that the computer uses to parse data, in order to find minute discernible patterns that become clearer with each parsing cycle; in essence, this is a probabilistic Bayesian machine that becomes more powerful as it processes more data. Because these methods require very large data sets to achieve a good deal of accuracy, they are a natural link in the big data chain. Cities, as enormous multidimensional data generators with complex problems, are clear grounds for experimentation.

The current and potential value of these technologies is enormous, and their criticality in driving the great diversity of data-driven applications that our modern societies live has led to enthusiastic adoption by citizens and cities worldwide. Given the ability of our current AI technologies "to make appropriate generalizations in a timely fashion based on limited data" (Kaplan, 2016, p. 6), they present new ways of understanding and proposing solutions to urban problems, as well as interventions that stimulate new ways to use the city. Cities move beyond the thousands of apps currently residing in our smartphones, capitalizing these algorithms to perform functions such as audio filtering for emergency and safety scenarios, vehicle detection and classification to automate traffic lights and optimize intersections, or facial recognition to personally identify an individual using video data (which can detect a potential terrorist in a crowd, or activate your smartphone to pay for dinner at a fast-food restaurant). In fact, many of the technologies that are poised to transform our cities in the coming future, from autonomous vehicles to smart grid systems, have a variety of AI systems at their core; here we must pause and consider the potential consequences of an algorithm-driven future. As Yuval Harari (2018, p. 470) suggests, the "coming technological revolution might establish the authority of Big Data algorithms, while undermining the very idea of individual freedom." Moreover, when corporations and governments are empowered by analytical tools and access to data (including people's personal data), they may make accurate predictions that seem way beyond what humans or social groups might achieve, and directing social behavior outside public forums.

There is a profound lack of synchronicity between the potential impact of AI technologies in our modern societies and our cultural discussions around them. While this partly stems from the fact that the field is quite technical and complex, the main cause is that most AI-related research simply remains in the hands of the researchers or private corporations. In fact, only 6% of presenters at major AI conferences share their code, which is also notably absent from top publications, such as *Nature* and *Science* (Hutson, 2018), in turn hindering the reproducibility of findings that would appear to be game changers in their fields. Furthermore, the values and impact of AI must be considered beyond its common fields of computer sciences and mathematics: the changes these technologies will bring are radical, and their ramifications will reach multiple aspects of social life.

For instance, Manyika et al. (2017) predicts that in less than 20 years, half of today's jobs could be carried out by AI-created algorithms. Industries like transportation and logistics, legal services, retail, and finance are being profoundly transformed in a clearly Schumpeterian fashion; for example, today 70% of financial transactions are already done algorithmically by computers, without direct human supervision (Helbing, 2019). These hyper-trading decisions are made in nanoseconds, far surpassing human biology, a distinct non-transferable advantage that computers have over humans. Here we must remember that these jobs do not take place in a theoretical space, but rather are actually performed by humans living in physical cities, in turn comprising a critical component of our social urban fabric. The current industrial configuration in our cities may feature high vulnerability, as in the case of Las Vegas, where according to a recent study over 60% of the current workforce faces the risk of losing jobs to automation in the near future (Chen, 2017). While machines replacing human labor is hardly a new phenomenon, it is the scale and fast pace of AI systems adoption that will put an unprecedented number of jobs at risk across a myriad of industries (Frey and Osborne, 2017) and strain our social capacity to positively absorb the technology.

Beyond the expected impact in the labor market, there are other considerations when designing and deploying large-scale AI systems. Their potential effects across our social systems are due in part to unprecedented decision-making authority, and the faith we humans vest in the accuracy of these technologies. While many of these decisions are of little consequence in our daily lives, some of them (such as the ones used in facial recognition technologies at security checkpoints or those used to score the social value of a person) can easily create personal and social havoc. In their individual capacity as well as in their compounded fashion, we humans have become over-reliant on them, hence the urgent need to recognize that many of these decisions are imperfect. Still, many algorithms in use today are just a reduction of much more complex realities. Even worse, some are potentially faulty due to inherent bias in the AI's decisions resulting from imperfect data used to train them; this has been demonstrated in human detection algorithms with greater rates of detection error for African-American people in relation to whites after minorities were under-represented in the training data sets (Buolamwini and Gebru, 2018). Thus, AI is "a cultural shift as much as a technical one" (Crawford and Calo, 2016).

Conclusions

How can the power of artificial intelligence be balanced to promote more efficient and stimulating urban experiences without losing the central human dimension of our cities?

Well-intended examples abound. Microsoft's chief environment scientist (Joppa, 2017) proposes creating a portfolio of open application programming interfaces (APIs) infused with AI that would allow people develop software to map and analyze environmental data. Vazifeh et al. (2018), in analyzing more than 150 million taxi trips in New York, demonstrated that it is possible to reduce the number of vehicles by more than 30%, transporting the same passengers by optimizing the chaining of these trips. By crunching data and testing multiple scenarios for different public policies, governments can nudge citizens to do what will eventually benefit more people.

What still amazes the scientific community and raises concerns among critics is that AI results are driven by the correlations AI algorithms find among a myriad of examples—which, as Kaplan (2016) puts it, "can yield insights and solve problems at a superhuman level, with no deeper understanding or causal knowledge about a domain." Therefore, although AI has been presenting better results, and indeed can usually outperform human decisions based on direct experiences or fewer examples, the fact remains that AI works as long as the world presents itself in a machine-readable data format. This risk is not new: the "metrics fixation" (Muller, 2018) is found throughout the history of science, and more recently has pervaded a variety of areas from public health to education and governance. It is essential to find good metrics for urban issues; but focusing only on issues that produce quantifiable data risks losing sight of other equally important problems. And because the generation and capture of quantifiable data greatly depends on technology, quantifiable metrics are more abundant in developed areas of the world. Would urban planners risk neglecting urban problems that afflict poor areas because technological and methodological tools intrinsically favor rich areas?

Despite growing international collaborations among researchers in developed and developing countries, data flows still follow South-North routes, while expertize moves in the opposite direction. Without similar technical and financial resources, researchers from the Global South seldom publish or file patents on data they help gather, jeopardizing capacity-building efforts of intended collaborations. As a result, scientific results seldom benefit the communities that generated or gathered the data. As Serwadda et al. (2018) point out regarding biomedical research, the point is not to stop data sharing, but rather to be aware of the "risk of imperiling rather than promoting public health" if a "colonial science" approach persists under the guise of open data. From a policy perspective, this creates a scenario in which public money from scant budgets in cities within developing countries are spent on open data programs and effectively leveraged as scientific research and commercial patents elsewhere.

We close this paper arguing that current data-driven methods powered by abundance of digital sensing technologies and data availability might profoundly change the way we understand cities. Similar to what is happening to other fields, from medicine and biology to law and archeology, what is necessary is a balance between enthusiasm, knowledge, skepticism, and social debate. Innovative methods recently developed under the umbrella of artificial intelligence have demonstrated results that were impossible to achieve with traditional scientific methods. These innovations should be accepted with enthusiasm among urban scholars, planners, and policy makers. Such enthusiasm will arise from knowledge: if some of these methods are new within the computer science community, the bar to understand them within the urban studies is even higher. The lack of understanding risk creating a negative reaction, and it should not. However, both enthusiasm and knowledge must be balance with some skepticism: are these findings really novel in comparison with what we have before?; even if novel in themselves, are they useful?; more importantly, are they addressing the most dramatic issues cities are and will be facing in the near future? Finally, encompassing enthusiasm, knowledge, and skepticism, a social debate about which data we are collecting, who has access to it, and what we are doing with the findings are paramount. Data-driven approaches to urban phenomena involves the deployment of sensors, often embedded in urban infrastructures and personal devices, which are amassing huge amounts of data about individuals or social interactions. Data, as well as technologies, are not neutral, they are political. Thus, scientific results are not neutral, but politically charged—and, therefore, from the start should be subjected to social debates.

The question then becomes: how can we make efficient and equitable use of the abundance of data which is increasingly generated and collected in cities around world, combined with epistemological, methodological, and technological tools, to foster design and policy innovations that can improve cities in both the developed and developing world?

Data availability

Data sharing is not applicable since no datasets were generated or analyzed in this study.

Received: 10 December 2018 Accepted: 24 April 2019
Published online: 28 May 2019

Notes

1 "Artificial Intelligence." *Merriam-Webster.com*. Merriam-Webster, (n.d.) Web. 6 Dec 2018.

References

Acuto M, Parnell S, Seto KC (2018) Building a global urban science. Nat Sustain 1 (1):2–4. https://doi.org/10.1038/s41893-017-0013-9

Acuto M (2018) Global science for city policy. Science 359(6372):165–166. https://doi.org/10.1126/science.aao2728

Ang L, Seng KP, Zungeru AM, Ijemaru GK(2017) Big sensor data systems for smart cities IEEE Internet Things J 4(5):1259–1271. https://doi.org/10.1109/jiot.2017.2695535

Anjomshoaa A, Duarte F, Rennings D, Matarazzo TJ, Desouza P, Ratti C (2018) City scanner: building and scheduling a mobile sensing platform for smart city services. IEEE Internet Things J 5(6):4567–4579. https://doi.org/10.1109/jiot.2018.2839058

Attard J, Orlandi F, Scerri S, Auer S (2015) A systematic review of open government data initiatives Govt Inform Quater 32(4):399–418. https://doi.org/10.1016/j.giq.2015.07.006

Batty M (2008) The size, scale, and shape of cities. Science 319(5864):769–771. https://doi.org/10.1126/science.1151419

Brenner N (2013) Theses on urbanization. Public Cult 25(169):85–114. https://doi.org/10.1215/08992363-1890477

Brenner N (2018) Debating planetary urbanization: for an engaged pluralism. Environ Plann D 026377581875751 36(3):570–590. https://doi.org/10.1177/0263775818757510

Bunnell T, Maringanti A (2010) Practising urban and regional research beyond metrocentricity Int J Urban Reg Res 34(2):415–420. https://doi.org/10.1111/j.1468-2427.2010.00988.x

Buolamwini J, Gebru T (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research 81:1–15, Conference on Fairness, Accountability, and Transparency

Calabrese F, Colonna M, Lovisolo P, Parata D, Ratti C (2011) Real-time urban monitoring using cell phones: a case study in Rome. IEEE Trans Intell Transp Syst 12(1):141–151. https://doi.org/10.1109/tits.2010.2074196

Chen J (2017) Future job automation to hit hardest in low wage metropolitan areas like Las Vegas, Orlando and Riverside-San Bernardino. Institute for Spatial Economic Analysis, Redlands, CA

Ching T, Ferreira J (2015) Smart cities: concepts, perceptions and lessons for planners. In Geertman S, Ferreira J, Jr., Goodspeed R, Stillwell J (eds) Lecture notes in geoinformation and cartography planning support systems and smart cities. Springer, New York, pp. 145–168

Crawford K, Calo R(2016) There is a blind spot in AI research Nature 538 (7625):311–313. https://doi.org/10.1038/538311a

Frey CB, Osborne MA(2017) The future of employment: How susceptible are jobs to computerisation? Technol Forecast Soc Change 114:254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Harari YN (2018) 21 lessons for the 21st century. Spiegel and Grau, New York

Helbing D (2019) Towards digital enlightenment: essays on the dark and light sides of the digital revolution. Springer Nature, New York

Hutson M (2018) Artificial intelligence faces reproducibility crisis. Science 359 (6377):725–726. https://doi.org/10.1126/science.359.6377.725

Joppa LN (2017) The case for technology investments in the environment. Nature 552(7685):325–328. https://doi.org/10.1038/d41586-017-08675-7

Kaplan J (2016) Artificial intelligence: what everyone needs to know. Oxford University Press, Oxford, UK

Kessler MD, Yerges-Armstrong L, Taub MA, Shetty AC, Maloney K, Jeng LJ, Ruczinski I, Levin AM, Williams LK, Beaty TH, Mathias RA (2016) Challenges and disparities in the application of personalized genomic medicine to populations with African ancestry. Nat Commun 7(1):12521. https://doi.org/10.1038/ncomms12521

Kühnert C, Helbing D, West GB (2006) Scaling laws in urban supply networks. Phys A 363(1):96–103. https://doi.org/10.1016/j.physa.2006.01.058

Lecun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444. https://doi.org/10.1038/nature14539

Lefebvre H (1970) La Révolution urbaine. Gallimard, Paris

Li R, Dong L, Zhang J, Wang X, Wang W, Di Z, Stanley HE (2017) Simple spatial scaling rules behind complex cities. Nat Commun 8(1):1841. https://doi.org/10.1038/s41467-017-01882-w

Manyika J, Chui M, Miremadi M, Bughin J, George K, Wilmott P, Dewhurst M (2017) A future that works: Automation, employment, and productivity. McKinsey Global Institutte, New York

Muller JZ (2018) The tyranny of metrics. Princeton University Press, Princeton, NJ

Santos M (1985) Espaco e metodo. Livraria, São Paulo

Serwadda D, Ndebele P, Grabowski MK, Bajunirwe F, Wanyenze RK (2018) Open data sharing and the Global South—Who benefits? Science 359 (6376):642–643. https://doi.org/10.1126/science.aap8395

Shannon C, Weaver W (1969) The mathematical theory of communication. University of Illinois Press, Champaign, IL

United Nations (2015) World population prospects. New York: United Nations

Vanky AP, Verma SK, Courtney TK, Santi P, Ratti C (2017) Effect of weather on pedestrian trip count and duration: City-scale evaluations using mobile phone application data. Prev Med Rep 8:30–37. https://doi.org/10.1016/j.pmedr.2017.07.002

Vazifeh MM, Santi P, Resta G, Strogatz SH, Ratti C (2018) Addressing the minimum fleet problem in on-demand urban mobility. Nature 557 (7706):534–538. https://doi.org/10.1038/s41586-018-0095-1

West GB (2017) Scale: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies. Penguin Press, London

Acknowledgements
The authors would like to thank CNPq-Brazil, and Dover Corporation, Teck, Lab Campus, Anas S.p.A., Cisco, SNCF Gares and Connexions, Brose, Allianz, UBER, Austrian Institute of Technology, Fraunhofer Institute, Kuwait-MIT Center for Natural Resources, SMART—Singapore-MIT Alliance for Research and Technology, AMS Institute, Shenzhen, Amsterdam, Victoria State Government, and all the members of the MIT Senseable City Lab Consortium for supporting this research.

Additional information
Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://www.nature.com/reprints

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.