CHALLENGES AND OPPORTUNITIES OF URBAN BIG-DATA FOR SUSTAINABLE DEVELOPMENT

Ali Kharrazi[a,b]
[a] Graduate School of Public Policy, The University of Tokyo, Tokyo, Japan. E-mail: ali@pp.u-tokyo.ac.jp
[b]Advanced Systems Analysis Group, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Abstract

Cities are the most critical arena for advancing sustainable development and there is a growing enthusiasm of and recognition for the potential of applying urban big-data towards solving urban sustainability challenges. However, while the promise of urban big-data is real, there is currently a wide gap between its full potential and its realization. This article offers a brief overview of different sources of urban big-data and gives examples of emerging applications for urban sustainability in the area of urban mobility and urban energy management. Furthermore, a discussion on the socio-technical challenges of urban big-data is presented. To advance the sustainability applications of urban big-data, policymakers, practitioners, and businesses should view the emergence of these innovations through a lens of 'hospitality'. Through this lens, cities may strengthen institutions to govern the development of urban big-data, enhance a grassroots culture for diffusing emerging sustainability innovations, and invest in university-led research and education.

Introduction

Cities, particularly highly dense urban areas, are at the nexus of economic and cultural development, environmental degradation, and are central to our visions of sustainable futures. Currently, half and by 2050 more than two-thirds of humanity will live in cities, generating the majority of economic activity, innovation and cultural advancement (United Nations, 2015). At the same time, cities will create environmental pollutants, consume higher amounts of energy and increase our vulnerability to natural hazards. Decoupling environmental impacts from economic development, improving energy efficiency, and reducing the risks and vulnerabilities in cities present highly complex challenges to policymakers, business leaders, and urban citizens. It is therefore with certainty that the battle for sustainable development will be fought, won or lost, in cities. The significant role of cities is reflected across the Sustainable Development Goals (SDGs) and specifically in SDG 11, which aims to "make cities and human settlements inclusive, safe, resilient and sustainable".

One of the most significant objectives of the SDGs is an emphasis on the collection, analysis, and integration of data relevant to sustainable development. Reflecting this importance, the United Nations has called for a 'data revolution' to develop objective targets and scientifically grounded indicators to monitor progress, implement strategies, allocate resources, and increase the accountability of stakeholders toward sustainable futures. Towards this end, many sustainability researchers and practitioners view urban big-data as having a unique potential to develop, experiment with, and advance sustainable development in cities. The predictive analysis associated with big-data innovations has the potential to empower people and change how urban residents interact with each other, their surrounding environment, and urban infrastructures. However, the technological affordances associated with such innovations are in their infancy and their challenges and opportunities need to be better clarified and understood.

What is urban big-data?

Urban big-data results from massive amounts of dynamic and static data from urban infrastructures, facilities, organizations, and individuals which have been collected by urban governments, public and private enterprises, individuals, and citizen-scientists using a new generation of information technologies. These emerging technologies include, for example, radio-frequency identification (RFIDs), Internet of Things (IoTs), smart grids, and other future applications of information technologies to urban management. Urban big-data can be categorized into five main types: sensor systems, user-generated content, administrative data, private sector transactions data, and data from arts and humanities collections (Thakuriah et al., 2017). A summary of each type and example of the each urban big-data category is provided in Table 1.

Sensor systems refer to urban sensors that can generate data on the supply, demand, availability, and or inoperability of urban infrastructure which urban citizens interact with on a daily basis. These include, for example, transport systems, buildings, water utilities, vehicle movement, street lighting, and other remote sensing applications. These sensor systems can potentially be designed to communicate and interact with one another and achieve a future where the Internet of Things (IoTs) dynamically manages urban functions.

User-generated content refers to data created or crowd-sourced by urban citizens. Through the use of sensors or social media, individual urban citizens generate social, economic, and cultural data. These can be focused on feedback in solving daily urban problems, e.g., user generated urban traffic reports, or more long-term urban planning issues, e.g., development of pedestrian and commercial zones. User generated data can provide a live and rich source of data on urban activities and community behaviors.

Table1: Established and emerging sources of urban big-data

Type	Example
Sensor systems (infrastructure-based or moving object sensors)	Environmental, water, transportation, building management sensor systems; connected systems; Internet of Things
User-Generated Content ("social" or "human" sensors)	Participatory sensing systems, citizen science projects, social media, web use, GPS, online social networks and other socially generated data
Administrative (governmental) data (open and confidential micro-data)	Open administrative data on transactions, taxes and revenue, payments and registrations; confidential person-level micro-data on employment, health, welfare payments, education records
Private sector data (customer and transactions records)	Customer transactions data from store cards and business records; fleet management systems; customer profile data from application forms; usage data from utilities and financial institutions; product purchases and terms of service agreements
Arts and humanities data	Repositories of text, images, sound recordings, linguistic data, film, art and material culture, and digital objects, and other media

Source: Thakuriah et al., 2017

Administrative data refers to daily micro-data collected by urban governments. This may include data for example on sales tax, income tax, building licenses, real-estate transactions, and vehicle registrations. These administrative micro-data are usually of high quality and conducive to public data innovations as they become increasingly available through open-data initiatives. Despite common challenges associated with data governance and privacy, many cities throughout the world are supportive of open-data initiatives.

Private sector data refers to customer transaction data generated by companies. This may include customer profiles, marketing data, and data on consumption trends. Such data can be beneficial for understanding and predicting consumption demands of key urban resources such as water and energy.

Arts and humanities data refers to highly unstructured data that portray urban life through artistic expressions. These data sources include, for example, art, film, media repositories, images, and texts. Such data not only enhances the creativity but also importantly increases the capacity of researchers and practitioners towards more constructive public stakeholder engagement and diffusion of innovations.

Applications of urban big-data for sustainability solutions

The emphasis of urban big-data is not only on the importance of size and variety of the data obtained, but on what new insights can be achieved and how it can enhance existing tools and indicators relevant to urban sustainability. While there is no doubt that we are witnessing a significant growth in volume, variety, velocity, and veracity of urban data, how such data can be interpreted to address and facilitate innovations towards solving many of our sustainability challenges is less clear. The emergence of big-data therefore needs to be acknowledged with caution and avoided to be overused as a buzzword without any clear direction for solving pressing wicked urban sustainability problems. Two areas which urban big-data has major potential to solve persistent urban challenges and lead to sustainable transformations are urban mobility and urban energy management.

Urban mobility

New urban transportation approaches resulting from urban big-data analytics that encourage ride sharing, route optimization, and on-demand vehicles can significantly clean up traffic congestion and result in lowering air and noise pollution levels. In this avenue, Seoul Metropolitan Government's effort to develop new transportation routes for the city's late night commuters is a good example. During late night and early morning hours, Seoul's citizens, especially low-income workers, had difficulty in finding reliable, safe, and inexpensive transportation to commute to work. As night bus services were not in operation and few metered taxis available, people were forced to use expensive and illegal makeshift taxis which would at times even refuse to take on the late-night passengers. In response, city officials decided to establish new night bus routes for the burgeoning metropolis. However, the city officials were not certain of the most effective and optimized routes that would cater to public demand. In search of a solution, city officials analyzed late-night mobile phone call usage patterns to better understand common departure and destination points across the city's urban hubs and outer districts (Seoul Metropolitan Government, 2017). This urban big-data approach allowed city planners to develop a route map that could optimally serve the city's late night commuters.

Traffic congestion is one of the costliest challenges endemic in all cities which drastically increase the wasting of fuel, loss of productive work time, and air pollution. Urban big-data is enabling cities to find new solutions to tackle the wicked problem of traffic congestion through ride-sharing. In New York City, new regulations require all taxis, including private for-hire vehicles such as Uber, Lyft, and other companies, to release all ride data to the city's databases (TLC, 2017). This comprehensive database includes pickup and drop off locations, duration, fares, and other details of each passenger trip within the city. This database will enable research on car-sharing approaches and transport optimization strategies based on the size, capacity, and travel time of the city's taxi fleet. Ride-sharing strategies can significantly reduce congestion, pollution, and fuel consumption. In this avenue, through the use of New York City's taxi transportation data, researchers have demonstrated that only 2,000 vehicles (only 15% of the city's taxi fleet) can potentially serve 98% of the city's taxi demand through ride-sharing and route optimization (Alonso-Mora et al., 2017). Such data-driven innovations are especially well suited to the emergence of future autonomous vehicles serving highly congested urban areas.

Urban energy management

Urban big-data can strengthen our capacity for sustainable transformations by increasing the ability to measure environmental flows and to employ environmental accounting of natural resources, energy, and pollutants. In highly dense urban regions, residential and office buildings are responsible for significant amounts of energy consumption and production of pollution. One of the most challenging aspects of improving the energy efficiency of a city has been the complex task of measuring energy use at the level of individual buildings. Towards this end, to gain more insight on how buildings are using energy and identify poorly performing buildings, cities around the world are legislating new data laws whereby building owners and managers are required to share their energy use data with city officials. New York City's Local Law 84 is one example of such legislation which currently requires large and from 2018 onwards mid-size buildings to evaluate and report both their energy and water usage (City of New York, 2009). By doing so, city officials can benchmark the energy efficiency of a city and compare buildings based on attributes such as age, size, and use type with their local peers and implement policies and incentives to maximize the building's energy efficiency with available technologies and practices. In the next step, cities are not only seeking more granular building energy use data, i.e., monthly, daily, and data from smaller size buildings, but also information on buildings assets. Building asset data refers to, for example, the age and model of the heating and cooling infrastructure of a building or insulation design and window configurations. These additional data can enhance the ability of city officials and energy saving businesses to target buildings with energy efficiency policies, practices, training, and technological diffusion strategies.

Challenges of urban big-data

The potential of urban big-data and its opportunities for policies and strategies relevant to sustainability is increasingly broadly recognized by both the public and private sectors. However, the analysis of data and its innovative applications have not yet reached a solid level of maturity. While the size and variety of data are increasing and urban policymakers and businesses are formulating strategies to support open-data initiatives; for many urban sustainability challenges, the data analysis and interpretation necessary for extracting actionable knowledge remain as clear bottlenecks. Furthermore, as the value of data significantly grows when it can be linked with other data, data-integration of the various types of urban big-data is still a major challenge. In this avenue, urban policymakers and businesses should be cautious in avoiding technological lock-ins and non-interoperability of urban big-data sources. Technological lock-in refers to a form of path dependence whereby a specific technological standard prevails and the city is locked-in to the standard even though new and better alternatives emerge in the market. The risk of technological lock-in is that a city may be locked-in to proprietary or black-box operating systems and lose control of critical urban data and information relevant to the management of urban infrastructures. Non-interoperability refers to the risk where an urban system's main infrastructure and databases are unable to communicate and exchange data with one another. The risk of non-interoperability is that a city would maintain, for example, different sensor networks solutions for its water, energy, and waste infrastructure, where each network could not communicate with the other in a standardized way. This would in effect reduce the value of the urban data generated by hindering critical data linkages. Therefore, urban policymakers and businesses should avoid creating dependence on very few technology providers and consider more support for open-source and inter-operable urban big-data practices.

In addition to technological challenges, the challenges of urban big-data from the viewpoint of public policy also need to be thoroughly considered. Given the infancy of urban big-data and its applications relevant to sustainability challenges, urban policymakers, practitioners, and businesses should approach urban big-data through what the late Claudio Ciborra termed as the lens of 'hospitality' (Ciborra, 2004). Through this lens, we are able to extend courtesy to the unknown and at times alien technological culture of urban big-data and therefore are able to implement its affordances and better understand its advantages and disadvantages. Towards this end, cities can be more hospitable, receptive, and open-minded to urban big-data and enhance its applications towards urban sustainability by establishing and strengthening mayoral institutions specific to the governance of urban big-data, enhancing a grassroots culture for applying big-data, and investing in university-led urban big-data research.

Developing urban institutions for the governance of urban big-data

At the heart of applying urban big-data towards sustainability are urban citizens and their daily life and business activities. This people-centric perspective requires urban policymakers to focus on what actors and stakeholders should be involved and what incentives and concerns would be important in the diffusion and application of big-data innovations. These issues are best approached by central mayoral or urban institutions which can govern, compile, legislate, standardize, and financially support urban big-data initiatives. In this avenue, New York City Mayor's Office of Data Analytics (MODA), London's City Data Team, and the position of Chief Technology Officer in Amsterdam are pioneering institutions in governing the emergence of urban big-data innovations for sustainability practices. An increasing need for mayoral institutions to govern urban big-data is not surprising; as with other urban utilities such as water, electricity and natural gas, data can also be viewed as a utility requiring public control and regulation.

Through such institutions, the collective hospitality of a city towards the emergence of urban big-data can better accommodate and respond to its many challenges. These challenges may include data ethics, privacy, and potential frictions that urban big-data innovations may have with traditional urban dynamics and business models. Through such institutions, a city can also better address the interoperability of various infrastructures and sensors and provide leadership in legislation concerning data access, ownership, and support for open-data initiatives. Open-data facilitates knowledge sharing and promotes citizen participation, transparency, the reduction of information inequality, and grassroots democratic engagement among urban citizens (The World Bank, 2015). Open data is especially beneficial towards decreasing the risk associated with data accessibility and affordability and for fostering innovative solutions to social, economic, and environmental challenges of urban sustainability.

Enhancing a grassroots culture for applying urban big-data

The planning, design, and diffusion of urban big-data are often viewed in the capacity and responsibility of large information technology corporations. However, a top-down approach to the development of urban big-data may not successfully engage urban citizens and may weaken data openness, transparency, and inclusiveness. Indeed, cities must strike a delicate balance between heavyweight technological corporations and the general public, urban communities, and technically-skilled activists. Without such a balance, cities risk a repeat of urban conflicts setting central urban planners like Robert Moses against community activists such as Jane Jacobs (Townsend, 2013). Cities are inherently organic ecologies and urban sustainability should be viewed as challenges of organized complexity. As we have witnessed before, rationalistic, central, and top-down urban planning approaches are at times in conflict with the organic qualities of a city. The risk of such conflicts can also be expected through the emergence of urban big-data innovations.

Through a grassroots approach, information asymmetries between large information technology corporations and local communities can be minimized and emerging innovations better contextualized within the urban communities. By empowering urban citizens and local community businesses, the risk of disadvantaging segments of society without the knowledge or skills sets necessary to benefit from urban big-data innovations can also be minimized. Inclusive and grassroots approaches for urban big-data allows for the maximization of the collective intelligence of society, the promotion of pluralism, and digital democracies (Helbing and Pournaras, 2015). By enhancing a grassroots culture and strengthening common skill sets necessary for leveraging big-data applications, urban citizens can begin to crowdsource innovations and measure their progress in tackling urban sustainability challenges.

Investing in urban big-data research and education

Universities and research institutions should lead the 'hospitality' towards urban big-data and use cities as both laboratories and classrooms in applying the emerging technologies towards solutions for urban sustainability challenges. A critical challenge for urban big-data continues to be its transdisciplinarity and the lack of people with specialized skills and focus in the area. Common skills and knowledge relevant for advancing big-data for urban sustainability are for example statistics, database engineering, data mining, geographic information systems (GIS), spatial analysis, programming, and information systems design and management. By investing in research and education and university-industry collaboration in urban big-data, a new generation of transdisciplinary researchers and practitioners can be trained to innovatively leverage data towards tackling urban sustainability problems. Towards this end, the emerging field of urban informatics is gaining importance among educational institutions around the world. Urban informatics is a transdisciplinary field which focuses on the broad overlap of people, places, and technologies (Goodspeed, 2017). Pioneering educational programs and research laboratories in this field are the Center for Urban Science and Progress at New York University; Amsterdam Institute for Advanced Metropolitan Solutions; The Bartlett Centre for Advanced Spatial Analysis at the University College London; Senseable City Lab at the Massachusetts Institute of Technology; Urban Informatics Research Lab at the Queensland University of Technology; and The Beijing City Lab. These institutes focus on interdisciplinary research on the science of cities and are well placed in training future urban big-data specialists and piloting solutions to the social, environmental, and economic problems of urbanizations.

Conclusions

Cities are the most important arena in humanity's battle for sustainable development and as emphasized in the 11th SDG, our visions of sustainable futures cannot be achieved without significantly transforming how urban spaces are managed. In this avenue, the emergence of big-data innovations presents a unique opportunity to leverage large and varied urban data towards better decision making and strategies for tackling urban sustainability challenges. While the five categories of urban big-data, i.e., sensor systems, user-generated content, administrative data, private sector transactions data, and arts and humanities data, continue to increase in quality and availability, the challenge remains to explore their collective synergies and applications towards urban sustainability challenges. Two areas with significant potential for channeling big-data towards urban sustainability solutions are in urban mobility and energy management. Through targeted legislation and practices cities can benefit from urban big-data analytics and tackle the problems of urban traffic congestion and excessive building energy use and pollution.

However, before the potential of urban big-data can be fully realized, many socio-technological challenges described in this article must be addressed. These challenges include not just the obvious technical data issues ranging from data acquisition to interpretation but issues of innovation management such as technological lock-ins and non-interoperability of urban big-data sources. Furthermore, the challenges of urban big-data from the viewpoint of public policy can be approached through a lens of 'hospitality', whereby policy-makers, practitioners, and businesses can be more receptive and open-minded towards the technology's emergence and its sustainability applications. Specifically, cities can strengthen institutions which can govern and guide the development of urban big-data, enhance a grassroots culture for applying and diffusing emerging sustainability innovations, and invest in university-led specialized research and education programs on urban big-data. Urban big-data has the potential to advance urban sustainability; however, big-data is not a magic bullet and it is imperative to encourage new perspectives on its challenges and opportunities towards the sustainable transformation of the management of cities.

References

- ✓Alonso-Mora, J., Samaranayake, S., Wallar, A., Frazzoli, E., & Rus, D. (2017). On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment. *Proceedings of the National Academy of Sciences, 114*(3), 462–467.
- ✓Ciborra, C. (2004). *The Labyrinths of Information: Challenging the Wisdom of Systems*. Oxford, UK,: Oxford University Press.
- ✓City of New York. (2009). Local Law 84 (LL84): *Benchmarking*. Retrieved from http://www.nyc.gov/html/planyc2030/downloads/pdf/ll84of2009_benchmarking.pdf
- ✓Goodspeed, R. (2017). Urban Informatics: Defining an Emerging Field. In L. A. Schintler & Z. Chen (Eds.), *Big Data for Regional Science*. Routledge.
- ✓Helbing, D., and Pournaras, E. (2015). Society: Build digital democracy. *Nature,* 527(7576), 33–34.
- ✓NYC Taxi and Limousine Commission (TLC). (2017). NYC Taxi and Limousine Commission (TLC) Trip Record Data. Retrieved from http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml
- ✓Seoul Metropolitan Government. (2017). *Night Bus (Owl Bus): Route Design Using Big Data*. Retrieved from https://www.seoulsolution.kr/en/node/2590
- ✓Thakuriah, P., Tilahun, N. Y., & Zellner, M. (2017). Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery. In P. Thakuriah, N. Y. Tilahun, & M. Zellner (Eds.), *Seeing Cities Through Big Data: Research, Methods and Applications in Urban Informatics* (pp. 11–45). Springer International Publishing Switzerland. https://doi.org/10.1007/978-3-319-40902-3_2
- ✓The World Bank. (2015). *Open Data for Sustainable Development*. Washington, DC. Retrieved from http://pubdocs.worldbank.org/en/999161440616941994/Open-Data-for-Sustainable-Development.pdf
- ✓Townsend, A. M. (2013). *Smart Cities: Big Data, Civic Hackers & Quest for a New Utopia*. New York, NY: WW Norton & Company.
- ✓United Nations. (2015). *World Urbanization Prospects: The 2014 Revision*. ■

Role of Big Data in SDG Implementation

The Center for Strategic and International Studies (CSIS) Project on Prosperity and Development and the Japan International Cooperation Agency (JICA) Research Institute have launched a report titled, 'Harnessing the Data Revolution to Achieve the Sustainable Development Goals: Enabling Frogs to Leap,' which defines the data revolution and identifies challenges and opportunities that it presents for implementing the SDGs. The report reviews the necessary elements for an enabling environment to leapfrog data technologies, presents country case studies, identifies key challenges in executing the sustainable development data agenda, and provides recommendations for how the international community can play a constructive role in the data revolution. Among the leapfrog data technologies that could support data collection and the formulation of policy based on data are satellite mapping, wearable technology, and cellular technology, such as mobile banking.

To play a constructive role in the data revolution, the report recommends that the international community: focus on the foundation necessary to facilitate leapfrogs around all types of data; increase funding for capacity building as part of an expansion of broader educational development priorities; highlight, share and support government-driven approaches to data; increase funding for the data revolution and coordinate donor efforts; coordinate UN data revolution-related activities with an expanded Global Partnership for Sustainable Development Data (GPSDD); and secure consensus on data sharing, ownership and privacy-related international standards.

For more information, access:

https://www.csis.org