Chapter 30

Introduction to Urban Big Data Infrastructure

Michael F. Goodchild

Rapid progress is being made in the development of infrastructure for handling urban big data, as will be evident from even the most cursory examination of the eight chapters in this section. Big data require the ability to handle unprecedented volumes of data, often in near-real time, and to fuse and conflate data from multiple sources with different degrees of quality. But in addition, the nature of infrastructure should be interpreted broadly, as encompassing not only data, but also the software needed to handle the data, the people who possess the requisite skills, and the decision-makers and general public who make use of the products of urban big data and may also contribute data through crowdsourcing. Moreover, no discussion of urban big data can escape the ethical issues that are raised by the technology and its use, especially the thorny issue of privacy. Urban big data infrastructure is clearly a vast topic, and these eight chapters can do no more than scratch the surface. The following paragraphs give a brief introduction to each chapter and explain how the various contributions fit together. At the end, a short discussion suggests some of the topics that might be covered in a longer review, and gives an overall assessment of this part of the book.

In Chap. 31, Ningchuan Xiao and Harvey Miller expand on the definition of urban big data, explaining its role in concepts of smart mobility, the smart city, and enhanced digital infrastructure. They review many sources of urban big data, from sensors to crowdsourcing, and argue strongly for open access as a key to supporting many potential applications. Some well-chosen stories are used to identify use cases, and the example of access to real-time data on transit vehicles is used to demonstrate some of the technical challenges.

While ethical issues are often regrettably left till last, we have chosen to raise questions of privacy early in the section. Chapter 32 by Jerome Dobson and William Herbert discusses geoprivacy, the threat to individual privacy that originates with

M. F. Goodchild (✉)
University of California, Santa Barbara, Santa Barbara, USA
e-mail: good@geog.ucsb.edu

the widespread capturing of an individual's coordinates, often without that individual's knowledge and conscious consent. Regulation varies from country to country and even within countries, and while the European Union has recently adopted comprehensive protection of user privacy, there has been little progress in the USA.

Accurate surveying of property has existed for centuries, but it has generally been assumed that a point can lie in at most one property. Today, this may no longer be true: In condominiums, for example, properties can be stacked on top of each other, requiring a three-dimensional (3D) approach. In Chap. 33, Lin Li provides an extensive review of the complex ownership geometries that can now be dealt with using three-dimensional techniques and digital representations.

Chapter 34 follows directly from Chap. 33 by providing a comprehensive review of techniques for 3D digital modeling of city structures. Much of this interest comes from the construction industry, whose building information modeling (BIM) provides techniques for capturing not only architectural plans, but also as-built information on building infrastructure and use. The chapter compares BIM with City Geography Markup Language (CityGML), a product of the geospatial community that brings spatial database modeling indoors, allowing a full integration between outdoor applications that are largely 2D, and indoor functions in full 3D.

The sequence of chapters on 3D representations of cities ends with Chap. 35, based on Esri's CityEngine. City planning requires consideration of buildings in context and specifically with the ways in which planners regulate the development of neighborhoods. CityEngine was developed as a multipurpose planning tool that is capable of implementing regulations, providing perspective visualizations of plans, and supporting many of the functions of city government. The chapter provides ample illustration of the applications of the software and its implications for geodesign and the planning process.

Today's cities are complex and growing more so as a result of recent investments in digital infrastructure. The massive volumes of data that are now available, and the speed at which decisions are needed, argue in many cases for the use of high-performance computing (HPC). Cyber geographic information systems (CyberGIS), the topic of Chap. 36, use HPC to address many such applications, extending conventional GIS to take advantage of massive computational and communication technologies.

Chapter 37 focuses on spatial search, the process that allows users to find and assess big data resources and judge their fitness for a given application. Techniques of spatial search became necessary beginning in the early 1990s, as the availability of geospatial data began to outstrip any user's knowledge of where to look. Data warehouses, geolibraries, and geoportals are all responses to the need to be systematic about the storage of geospatial data. The chapter reviews the relevant techniques, including the concept of metadata, that is, data that allow a user to assess the fitness of a given data set.

Finally, Chap. 38 addresses the Internet of things (IoT), a term that describes sensors of various kinds that are connected to the Internet. Sensors might be fixed in space, such as closed-circuit television (CCTV) cameras, carried on vehicles, or

carried by humans, often in the form of smartphone functions. IoT is clearly an important aspect of the smart city and of urban big data.

Big data infrastructure is a means to an end, rather than an end in itself. While Part IV has provided an overview of some of the foundational issues, the reader will have to look further for a complete view of the role of this infrastructure in enabling the functions of the modern city. Some of that can be found in other sections of this volume, and some is surely yet to emerge. While we can perhaps see and share some of the excitement over IoT or CityEngine, the eventual value of these tools is still difficult to predict. There is a “build it and they will come” sense to big data infrastructure, but also a sense that some of the eventual outcomes are unanticipated and may well have costs that exceed their benefits. Chapter 32 on privacy is perhaps a foretaste of what may arise as the technologies of surveillance proliferate.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.