Chapter 29

User-Generated Content and Its Applications in Urban Studies

Wei Tu, Qingquan Li, Yatao Zhang, and Yang Yue

Abstract The emergence of Web 2.0 and mobile Internet produces massive user-generated content (UGC), including geo-tagged photos, social network posts, street view images, and crowdsourced GPS trajectories. UGC creates unprecedented opportunities to sense what was previously hidden in the physical surfaces of cities and to portray the interactions of infrastructures, geo-information, and people; therefore, it is not only a new lens for urban space but also leads to innovative applications. In this chapter, we will introduce several typical types of UGC, such as geo-tagged photos, social media data, crowdsourcing GPS trajectories, and videos. We showcase ways in which user-generated big data can be harvested and analyzed to generate invisible and impressionistic landscapes of urban dynamics and to stimulate innovative applications. We discuss typical UGC-driven applications to demonstrate the potential of UGC in revealing how urban spaces are perceived by the public, establishing links between tangible artifacts and physical-cyber-social spaces. This fosters alternative approaches to urban informatics that better capture the intricate nature of urban space and its dynamics.

W. Tu · Q. Li (✉) · Y. Yue

Guangdong Key Laboratory of Urban Informatics, Key Laboratory for Geo-Environmental Monitoring of Coastal Zone of Ministry of Natural Resource, and Shenzhen Key Laboratory of Spatial Smart Sensing and Service, Shenzhen University, Shenzhen, China
e-mail: liqq@szu.edu.cn

W. Tu

e-mail: tuwei@szu.edu.cn

Y. Yue

e-mail: yangyue@szu.edu.cn

Department of Urban Informatics, School of Architecture and Urban Planning, Shenzhen University, Shenzhen, China

Q. Li · Y. Zhang

State Key Lab of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China

e-mail: yatao@foxmail.com

© The Author(s) 2021

W. Shi et al. (eds.), *Urban Informatics*, The Urban Book Series,
https://doi.org/10.1007/978-981-15-8983-6_29

29.1 Introduction

Cities are the living spaces of more than 50% of the global population but occupy less than 2% of the Earth's land surface. Although the past decades have witnessed advances in the economy, the environment, and human health in urban areas, especially in developing countries, cities are still facing great challenges on the way toward a sustainable future. These challenges include traffic congestion, environmental pollution, waste management, vitality loss, and social inequality. Since 2000, the boom of information and communication technologies (ICT), Internet, and artificial intelligence (AI) has produced massive urban data. Therefore, urban studies are increasingly adopting an information-centric approach where they meet geographic information science (GIS), computer science, urban planning, etc. (Batty 2013; Li 2017).

When enabled with Web 2.0, mobile Internet, and smartphones, humans become sensors to perceive their immediate surroundings and thus produce multi-source and heterogeneous content, such as text, images, videos, and audio, that is, user-generated content (UGC) (Koskinen 2003; Wang et al. 2014). UGC denotes content that has been posted by users on online platforms, including Internet forums, blogs, wikis, Instagram, YouTube, Douyin, and social networks such as Weibo, Facebook, and Twitter (Cha et al. 2007; George and Scerri 2007; Goodchild 2007; Krumm et al. 2008; Lenders et al. 2008; Hollenstein and Purves 2010; Heipke 2010). The use of UGC has grown rapidly in recent years, because of its comparatively low cost, high penetration, and fast update. For instance, the popular Wikipedia (Fig. 29.1a), edited by worldwide volunteers, has become the largest encyclopedia in the world and continues to be updated following advances in science, technology, and society. Another example is OpenStreetMap (OSM; Haklay and Weber 2008; Fig. 29.1b) which attracts large numbers of volunteers who use GPS and fine-resolution imagery to produce a comprehensive base map covering 80% of all roads (Barrington-Leigh

Fig. 29.1 Representative user-generated content Web sites. **a** Wikipedia (<https://www.wikipedia.org/>); **b** OpenStreetMap in Shenzhen (<https://www.openstreetmap.org/#map=11/22.5322/114.0912&layers=T>)

and Millard-Ball 2017). Nowadays, OSM not only supports route planning and navigation services but also provides benefits to city planners with newly available urban data.

Classic urban studies generally rely on census data or field survey, which is expensive, labor-intensive, and of low temporal resolution. UGC enables urban studies to dive into the wave of big data (Aguilera et al. 2016). In general, UGC is produced by volunteers and thus contains volunteers' perceptions, preferences, or opinions about places, topics, and people. Accordingly, massive UGC provides unprecedented data sources for urban researchers to extra urban knowledge. On the other hand, UGC also motivates an alternative approach for conceptualizing and portraying the dynamics, structures, and characteristics of city. Consequently, UGC stimulates innovative urban applications which sense infrastructures, spaces, and people at all scales, reveals hidden urban knowledge, and makes real-time responses in support of urban emergency and long-term urban policies. Here, we sketch several types of UGC and their potential in urban sectors. The general framework of UGC-driven urban studies and insightful urban applications is reviewed. We discuss the challenges and future directions, including data quality and privacy, multi-source data fusion, integration of urban sensing, and urban governance.

The remainder of this chapter is organized as follows: Sect. 29.2 introduces four representative types of UGC, including geo-tagged photos, social media data, crowdsourcing GPS trajectories, and videos. Section 29.3 presents the general framework of UGC-driven urban studies and reviews typical urban applications. Section 29.4 discusses challenges and future directions. Section 29.5 concludes the chapter and discusses future work.

29.2 User-Generated Content

User-generated content has had a great impact on information-centric urban studies because of its appealing characteristics that crystallize the relationship between urban spaces and human activities with massive crowdsourcing data (Crooks et al. 2016; Jenkins et al. 2016; Thakuriah et al. 2016; Valdez et al. 2018). Accordingly, the sources and types of UGC are various (Heipke 2010; Mart et al. 2019; See et al. 2019). The focus here mainly concentrates on geo-tagged user-generated content as it provides opportunities to expose the hidden social, economic, and demographic information in urban spaces (Jenkins et al. 2016), which greatly benefits our understanding of the diversity of urban spaces and the complexity of urban dynamics. This section reviews several popular types of UGC and their characteristics, to provide a global overview of UGC, including geo-tagged photos, social media data, crowdsourcing GPS trajectories, and videos.

29.2.1 *Geo-Tagged Photos*

Geo-tagged photos are images uploaded to Internet forums and social networks by users. Usually, these photos are tagged with either explicit geographic coordinates or implicit forms of geo-information (e.g. point of interest or place name). There are two popular types of geo-tagged photos. One is sourced from the photo-sharing services, such as Flickr or Picasa, which allow users to share geo-tagged photos with text tags (Chen et al. 2018). Nowadays, there are many publicly available geo-tagged photos. For example, Yahoo Research Lab (Thomee et al. 2016) published one Flickr dataset YFCC100M containing 100 million images (<https://webscope.sandbox.yahoo.com/catalog.php?datatype=i>) for benchmarking purposes. MIT CSAIL (Zhou et al. 2018) published the dataset Place including 10 million photos of urban landmarks (<http://places2.csail.mit.edu/>). These photos, coordinates, and timestamps can be used to generate user footprints (Alivand and Hochmair 2017). Meanwhile, tagged texts provide auxiliary information with certain models, e.g. topic probability models. Through extracting the information hidden in these photos, researchers can effectively detect the temporal activities of photo takers and further analyze the behavior patterns of urban citizens.

Another type of geo-tagged photo is sourced from street view images collected by vehicles or volunteers, such as Google Street View (Hara et al. 2013; Li et al. 2015). Street view images usually contain one panoramic image and the corresponding location and therefore provide a sequence of images along a road. Different from remote sensing images monitoring geographic objects from above (aerial or space), the major advantage of street view images is the access they provide to urban landscapes from a pedestrian-like angle (Li et al. 2015; Cao et al. 2018). Consequently, street view images have had a significant impact on street level research, on such topics as urban greenery (Li et al. 2015), sidewalk accessibility (Hara et al. 2013), and the demographics of neighborhoods (Geburu et al. 2017).

Using innovative technologies such as computer vision and semantic annotations, geo-tagged photos have been used to extract massive knowledge about urban places and human beings. With regard to urban places, geo-tagged photos enable us to assess urban landscapes (Geburu et al. 2017; Li et al. 2015), including, for example, the distribution of urban infrastructure. In terms of human beings, they offer an opportunity to explore human social and mobility patterns at multiple geographic scales (Alivand and Hochmair 2017; Zhang et al. 2018). Furthermore, researchers can leverage them as a lens to articulate the relationship between urban spaces and human beings.

29.2.2 *Social Media Data*

Social media data contribute another valuable form of content to urban studies, especially location-based social networks (LBSN) (Kim et al. 2017; Shelton et al. 2015;

Thakuriah et al. 2016). In 2018, there were over 3 billion active social media users, and almost 3 billion active users of mobile social media (Mart et al. 2019). Generally, LBSN data provide various perspectives on social, economic, and demographic aspects in urban spaces. Through embedding social media data into urban spaces, the link to human beings is established, enabling the tangible and comprehensive understanding of human–environment interactions (Mart et al. 2019).

To date, there have been many substantial studies using LBSN data (e.g. Foursquare, Twitter, Airbnb, and Weibo) to portray urban dynamics. Table 29.1 lists the publicly available social media content. Foursquare data usually include place information, including check-ins, ratings, tips, and photos. Foursquare data have been used to identify users’ perceptions and preferences in urban spaces through the identification of the most visited or checked-in places (Agryzkov et al. 2016; Mart et al. 2017). Twitter and Weibo are other commonly used social media datasets. The coordinates and timestamps associated with social media content of Twitter can be used to detect the spatiotemporal patterns in people’s presence and activities (Crooks et al. 2015). Combined with natural language processing (NLP), Twitter is capable of detecting certain events, hot topics, culture distribution, urban functions, etc. (Yang et al. 2015; Tu et al. 2017; Tu et al. 2018a). Different from Twitter data, the content of Instagram is more visually related about the observed entity rather than text related, in the format of coordinates, photos, and corresponding descriptions (Giridhar et al. 2017). Thus, Instagram-based studies focus on the descriptions of a place through keywords and the activities happening in a place (Mart et al. 2019). Airbnb, one Web site offering information about temporal accommodation plays an important role in urban studies about rental homes. Meanwhile, Airbnb content provides an insight to observe tourism, especially in tourist cities.

Table 29.1 Publicly available social media data

Social media data	Description	Web link
Global Foursquare check-in dataset (Yang et al. 2015)	Contains 33,278,683 check-ins by 266,909 users on 3,680,126 venues (in 415 cities in 77 countries)	https://sites.google.com/site/yangdingqi/home/foursquare-dataset
Twitter dataset (Yang and Leskovec 2011)	Includes 467 million Twitter posts from 20 million users covering a 7-month period from June 1, 2009, to December 31, 2009	https://snap.stanford.edu/data/twitter7.html
Instagram dataset (Ferrara et al. 2014)	Contains information from 45,000 users of Instagram during the period from Jan 20 to Feb 17, 2014	http://www.emilio.ferrara.name/datasets/
Airbnb dataset	Contains reviews, listings, and neighborhood information in worldwide cities	http://insideairbnb.com/get-the-data.html

29.2.3 *Crowdsourcing GPS Trajectories*

The availability of crowdsourcing technologies facilitates the emergence and effective usage of geospatial data, which is of profound significance in the planning and management of urban spaces (Crooks et al. 2015; Jenkins et al. 2016). Crowdsourced GPS trajectories are usually collected by volunteers without professional services (Heipke 2010), implementing the concept of citizens as sensors proposed by Goodchild (See et al. 2019). So far, there have been many projects about crowdsourcing geospatial data (Heipke 2010), such as OpenStreetMap (OSM) (Planet 2019), Wikimapia, or HD Traffic™. OSM is probably the most prominent among all the crowdsourcing projects (Heipke 2010). The purpose of OSM is to establish a free, editable map across the world, supported by volunteers acting as sensors to collect geographic data (Barron et al. 2014). OSM has been widely used in a broad range of urban applications, from navigation to routing, from urban block division to urban function recognition (Crooks et al. 2015). In addition, digital footprints extracted from crowdsourced GPS trajectories are also important proxies. Digital footprints through time provide an insight to understand human mobility patterns and also offer access to the dynamic cognition of urban places.

29.2.4 *Videos*

Videos contain amounts of dynamic information about described phenomena and can greatly assist urban planning and management, such as urban scene understanding (Cordts et al. 2016), human activity analysis (Zhu et al. 2017), transportation surveillance (Chen et al. 2016), and emergency management (Schnebele et al. 2015). There are many ways to obtain video datasets, such as from YouTube videos (Douyin and Kuaishou), from social media platforms, urban surveillance videos, and street videos. Unlike the above three kinds of UGC data, although information in videos is wealthy and dynamic, it is relatively difficult to process videos quickly and efficiently due to their volume, noise, and diversity (Zhu et al. 2017). Lots of techniques for motion estimation, tracking, segmentation, and video filtering have been developed (Tekalp 2015). Nowadays, human activity and perception have become hot topics in urban studies. Videos from social media platforms, such as YouTube, can be utilized to perform spatiotemporal mapping of human activity, in the form of human activity recognition, sport mapping, weather impacts on human activities, crime detection, etc. (Zhu et al. 2017). Moreover, videos can reveal functions in urban scene understanding, such as those revealed by the Cityscapes dataset (<https://www.cityscapes-dataset.com/>; Cordts et al. 2016). This dataset provides a detailed annotated class list of urban stereo videos covering fifty cities, which can be used in semantic understanding of urban scenes (Cordts et al. 2016).

29.3 Urban Studies Driven by User-Generated Content

User-generated content contains massive hidden information, such as the users' socioeconomic status, preferences, opinions, and activity-mobility patterns (Jenkins et al. 2016; Mart et al. 2019; Thakuriah et al. 2016; Venerandi et al. 2015). Large volume UGC is stored, cleaned, and extracted to learn about phenomena in urban spaces and the interactions between urban functions and people. Consequently, UGC has been widely applied in urban studies, such as in urban planning, urban transportation, urban environment, and health. This section presents the general framework of UGC-driven urban studies and reviews representative urban applications.

29.3.1 *Framework for UGC-Driven Urban Studies*

Acquisition, integration, and analysis of UGC can be used to tackle the major issues that cities face, e.g. traffic congestion, urban growth, air pollution, public health, and urban safety. Generally, the framework of UGC-driven urban studies contains four layers from the bottom to the top as shown in Fig. 29.2: UGC harvesting, UGC management, UGC analytics, and smart urban applications.

In the UGC harvesting layer, single- or multi-source UGC is acquired from an online forum, vertical Web sites, and social networks. For example, posted Twitter messages about a city will be crawled for future data processing and analytics. In the second UGC management layer, the acquired UGC will be organized by locations, by users, or by associated topics. High-performance computing architectures and effective indexing structures that simultaneously incorporate spatiotemporal information, and texts will be built for efficient data manipulation. In the UGC analytics step, data mining (clustering and classification), and machine learning (e.g. logistics regression, decision tree, random forest, and support vector machine), deep learning (e.g. convolutional neural networks, deep residual networks, generative adversarial networks), and visualization will be used to recognize objects, patterns, and associations, and to speculate about causes and effects. In the smart urban application step, this extracted urban knowledge will be utilized by urban planners, transportation officials, environmentalists, and medical departments. In addition, the information will be disseminated to related people and organizations to improve urban living.

29.3.2 *Urban Planning*

Urban planning refers to social, economic, and political activities concerning the interconnectedness and complexity of urban spaces (Levy 2016). Urban planning is close related to many interactions of places and people, including urban form, land-use planning, locating transportation infrastructures, and designing urban interfaces.

Fig. 29.2 General framework of urban studies using user-generated content

UGC not only provides rich representations about urban space, but also opens access to human activity research (Crooks et al. 2015; Li et al. 2017; Longley and Adnan 2016).

The focus here mainly lies on two parts, namely human activity, and urban form and function. Regarding human activity, social media data collected from a great number of users, such as Foursquare, Instagram, or Twitter, provide the detailed descriptions of human activities within urban spaces (Mart et al. 2019), with which researchers can recognize activity patterns at suitable spatiotemporal scales. Recently, Tu et al. (2018a) fused large volume social media check-in data and mobile phone positioning data to extract city-wide human activities and portray their diurnal patterns. Gebru et al. (2017) inferred demographic information at neighborhoods across the USA from massive street view images. Studies of urban form and function address the aggregation of the physical shapes of urban spaces and the human activities that happen in these spaces respectively (Crooks et al. 2016). UGC provides large amounts of information that can be used to understand urban form and function and highlights how they influence each other (Crooks et al. 2015). Street network maps of OSM give detailed insights into urban form and are of fundamental importance in a range of applications. Other types of UGC, such as geo-tagged photos and social-media data, can be used to understand urban function (Gebru et al. 2017; Li et al. 2015; Cao et al. 2018). For example, Zhong et al. (2018) presented a

tweet-topic-function-structure framework to reveal spatial patterns from individual tweets. Their results demonstrated that when aggregating tweets by zones, the areas with the same topics form spatial clusters but of entangled urban functions. Using massive street view images, Zhang et al. (2018) developed a data-driven deep learning approach to map the distribution of city-wide human perception (e.g. safe, lively, beautiful, wealthy, depressing, or boring), which suggest the potential of massive UGC.

29.3.3 Urban Transportation

Transportation is essential to daily movements in the city. Quantities of urban-sensed data have been used to resolve problems in urban transportation and to build intelligent transportation systems (ITS; Wang et al. 2016). The social media platforms, mobile phones, and surveillance videos make it possible to generate rich social signals in a real-time manner and establish a data foundation for social transportation research (Zheng et al. 2016). UGC-based ITS can make use of various crowdsourced social signals to understand the social needs of transportation and combine needs and services to improve efficiency and effectiveness and make traffic conditions and citizen travel more convenient (Wang et al. 2016; Tu et al. 2019).

UGC can be used in a range of applications in urban transportation, for example, in mapping road networks, monitoring real-time traffic, or recommending travel routes. In terms of traffic monitoring, information obtained from social media platforms, such as Twitter, YouTube, and Flickr, encourages people to participate effectively in traffic tasks, such as identifying road hazards, and greatly cuts down on the related financial burden of government (Santani et al. 2015). In traffic management, social media data support shortest path computing, travel recommendation, etc., and can be improved by exploiting the content hidden in UGC (Wang et al. 2016). With respect to future green transportation, UGC that connects vehicles, people, and urban infrastructures can help to advance the efficiency of entire transportation systems and to promote reductions in fuel consumption and carbon emission (Wang et al. 2016).

29.3.4 Urban Environments and Health

The urban environment has a close relationship to the quality of human life and health, both of which should be emphasized in urban governance. The knowledge mined from social media data, mobile phones, and other UGC can provide opportunities to quantify aspects of the urban environment, such as urban green space (Li et al. 2015), air quality (Jiang et al. 2015), soundscapes (Aiello et al. 2016), and heat distribution (Overeem et al. 2013). Thus, fine-resolution maps of these environmental factors can help urban planners to improve residents' quality of life, surroundings, and health. For example, utilizing the green index of street view images, Li et al. (2015) assess

street-level urban greenery and provide suggestions for urban planners to reasonably improve the distribution of urban green spaces. Jiang et al. (2015) analyzed the spatiotemporal tendency in social media data using Sina Weibo (Chinese Twitter) in an effort to monitor air quality dynamically in large cities. Also, maps can be drawn by establishing a relationship between human perceptions and soundscapes extracted from social media (Aiello et al. 2016). In addition, smartphone battery temperatures can be used to estimate urban daily mean air temperatures by utilizing a heat transfer model in real time (Overeem et al. 2013).

29.3.5 Urban Safety

Citizens residing in urban areas may face fires, storms, heavy rainfall, traffic jams, and other hazards, which affect urban safety and human life. Therefore, it is important to detect urban emergency events in real time (Xu et al. 2016). Lots of messages from UGC, such as social media, volunteered photos, and videos, contain information about urban events and are important data sources to derive emergency events, capture their physical and social features, and help urban management departments to react quickly (Schnebele et al. 2015; Xu et al. 2016). Thus, event detection becomes a crucial issue in urban emergency management. There have been many studies focused on urban event detection. For example, some studies proposed adaptive algorithms to detect urban events through geo-tagged data from photo-sharing services (Papadopoulos et al. 2010). Making use of crowdsourcing to build an emergency management system is another choice (Oliveira et al. 2017). In addition, in order to detect emergency events in real time, the 5 W (What, Where, When, Who, and Why) characteristics are proposed to depict the spatial and temporal information of social media and thus to achieve detection goals (Xu et al. 2016).

29.4 Challenges and Future Directions

Recent UGC research has made great advances in the domain of urban studies. Many innovative urban applications have stimulated thinking about better urban living. Because of the complexity of cities (Batty 2007), this research presents various challenges to information-centric cities.

29.4.1 Data Quality and Privacy

Recently, with the growing interest in artificial intelligence, it has become possible to produce UGC not only by people but also by machines. Several studies have reported that fake messages are posted on Twitter (Fourney et al. 2017). Many

machine accounts have been created to disseminate special texts and photos with the objective of influencing specific groups of people. Consequently, UGC may be biased. When conducting UGC-driven urban studies, attention should be paid to the data quality issue to strengthen the reliability of the findings (Tu et al. 2018b; Jiang et al. 2019).

The privacy of UGC is another important issue. Scientific ethics should be highlighted for UGC research. Recently, a new General Data Protection Regulation (GDPR) was adopted in Europe and is likely to fundamentally reshape the way in which data are handled across every sector. The general public, Internet giants, and scientific communities should find an appropriate consensus on the collection, processing, and study of UGC.

29.4.2 Multi-source UGC Fusion

When thousands and even millions of users contribute to UGC, the results are often highly fragmented. For example, because most geo-tagged photos are shared by users with smartphones, the perceptions and preferences of people without smartphones cannot be captured. Tweets posted in tourist destinations and at landmarks tend to emphasize certain topics and opinions, resulting in bias with respect to the general population (Longley and Adnan 2016). Thus, careful selection of data sources is crucial if the reconstructed urban knowledge is to be complete and accurate. The results from a single source of UGC may be biased and contain only a part of urban knowledge. The misuse of UGC may consequently generate biased understanding. Fusion of multiple sources may be required to deepen our understanding of objects, people, and places in the city (Li et al. 2017). By integrating traditional urban data and alternative UGC, more and more comprehensive and wide-coverage urban solutions would be supported (Estima and Painho 2016).

29.4.3 Integrating Urban Sensing and Urban Governance

UGC can provide alternative data sources to sense the invisible city under the physical surface, for example, regarding urban deprivation (Venerandi et al. 2015), human mobility (Yang, Qu, Yang et al. 2019; Xu et al. 2019), urban areas of interest (Chen et al. 2018), urban vibrancy (Huang et al. 2019), and urban functions (Tu et al. 2017, 2018a; Zhong et al. 2018). UGC enables us to assess new dimensions of the city and to deepen our understanding of complex cities. However, these novel urban-sensing studies have not been well integrated with urban governance. How to take the sensed urban information into the workflow of urban governance is still an open question. UGC-driven urban policy-making will be necessary if we are to explore a new framework linking UGC to urban operation.

29.5 Conclusion

The prevalence of UGC provides an alternative data source for urban studies because of its characteristics of low cost, high penetration, and wide coverage. Massive UGC can not only sense invisible urban spaces but also provide fertile soil for breeding innovative applications. This chapter has summarized the four representative types of UGC: geo-tagged photos, social-media data, crowdsourced GPS trajectories, and videos. The general framework of UGC-driven urban studies has been presented, and smart UGC-driven applications in the city have been reviewed. The challenges and opportunities of UGC in urban studies have also been discussed, in order to provide insights for future urban informatics approaches. This will lead to the emergence of alternative urban informatics approaches that better capture the intricate nature of urban spaces and their dynamics.

Acknowledgements This research was jointly supported by the Natural Science Foundation of China (71961137003, 41401444), Shenzhen Scientific Research and Development Funding Program (#JCYJ20180305125113883), and China Scholarship Council (201708440434).

References

- Agryzkov T, Martí P, Nolasco-Cirugeda A, Serrano-Estrada L, Tortosa L, Vicent JF (2016) Analysing successful public spaces in an urban street network using data from the social networks Foursquare and Twitter. *Appl Netw Sci* 1(1):12
- Aguilera U, López-de-Ipiña D, Pérez J (2016) Collaboration-centred cities through urban apps based on open and user-generated data. *Sensors* 16(7):1022
- Aiello LM, Schifanella R, Quercia D, Aletta F (2016) Chatty maps: constructing sound maps of urban areas from social media data. *R Soc Open Sci* 3(3):150690
- Alivand M, Hochmair HH (2017) Spatiotemporal analysis of photo contribution patterns to Panoramio and Flickr. *Cartography and Geographic Information Science* 44(2):170–184
- Barrington-Leigh C, Millard-Ball A (2017) The world's user-generated road map is more than 80% complete. *PLoS ONE* 12(8):e180698
- Barron C, Neis P, Zipf A (2014) A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Trans GIS* 18(6):877–895
- Batty M (2007) Complexity in city systems: understanding, evolution, and design. In: de Roo G, Silva EA (eds) *A planner's encounter with complexity*, pp 99–122
- Batty M (2013) Big Data, smart cities and city planning. *Dialogues Hum Geogr* 3(3):274–279
- Cao R, Zhu J, Tu W, Li Q, Cao J, Liu B, Zhang Q, Qiu G (2018) Integrating aerial and street view images for urban land use classification. *Remote Sens* 10(10):1553
- Cha M, Kwak H, Rodriguez P, Ahn YY, Moon S (2007) I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system. In: *Proceedings of the 7th ACM SIGCOMM conference on internet measurement*, San Diego, California, USA, pp 1–14
- Chen M, Arribas-Bel D, Singleton A (2018) Understanding the dynamics of urban areas of interest through volunteered geographic information. *J Geogr Syst* 21(1):89–109
- Chen N, Chen Y, You Y, Ling H, Liang P, Zimmermann R (2016) Dynamic urban surveillance video stream processing using fog computing. In: *IEEE second international conference on multimedia big data (BigMM)* pp 105–112

- Crooks A, Pfoser D, Jenkins A, Croitoru A, Stefanidis A, Smith D, Karagiorgou S, Efentakis A, Lamprianidis G (2015) Crowdsourcing urban form and function. *Int J Geogr Inf Sci* 29(5):720–741
- Crooks AT, Croitoru A, Jenkins A, Mahabir R, Agouris P, Stefanidis A (2016) User-generated Big Data and urban morphology. *Built Environ* 42(3):396–414
- Cordts M, Omran M, Ramos S, Rehfeld T, Enzweiler M, Benenson R, Franke U, Roth S, Schiele B (2016) The cityscapes dataset for semantic urban scene understanding. In: *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)* pp 3213–3223
- Estima J, Painho M (2016) User generated spatial content-integrator: Conceptual model to integrate data from diverse sources of user generated spatial content. *ISPRS Int J Geo-Inf* 5(10):183
- Ferrara E, Interdonato E, Tagarelli A (2014) Online popularity and topical interests through the lens of Instagram. In: *Proceedings of the 25th ACM conference on hypertext and social media, Santiago, Chile, 1–4 September 2014*
- Fourney A, Racz MZ, Ranade G, Mobius M, Horvitz E (2017) Geographic and temporal trends in fake news consumption during the 2016 US Presidential Election. In: *Proceedings of the 2017 ACM conference on information and knowledge management, New York*, pp 2071–2074
- Gebru T, Krause J, Wang Y, Chen D, Deng J, Aiden EL, Fei-Fei L (2017) Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. *Proc Natl Acad Sci* 114(50):13108–13113
- George C, Scerri J (2007) Web 2.0 and user-generated content: legal challenges in the new frontier. *J Inf Law Technol* 2. SSRN: <https://ssrn.com/abstract=1290715>
- Giridhar P, Wang S, Abdelzaher T, Ganti R, Kaplan L, George J (2017) On localizing urban events with Instagram. In: *IEEE INFOCOM*, pp 1–9
- Goodchild MF (2007) Citizens as sensors: Web 2.0 and the volunteering of geographic information. *Int J Spat Data Infrastruct Res* 7:24–32
- Haklay M, Weber P (2008) Open street map: User-generated street maps. *IEEE Pervasive Comput* 7(4):12–18
- Hara K, Le V, Froehlich J (2013) Combining crowdsourcing and Google street view to identify street-level accessibility problems. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, pp 631–640
- Heipke C (2010) Crowdsourcing geospatial data. *ISPRS Journal of Photogrammetry and Remote Sensing* 65(6):550–557
- Hollenstein L (2010) Purves RJ (2010) Exploring place through user-generated content: Using Flickr tags to describe city cores. *J Spat Inf Sci* 1:21–48
- Huang B, Yulun Z, Zhigang L, Song Y, Cai J, Tu W (2019) Evaluating and characterizing urban vibrancy using spatial Big Data: Shanghai as a case study. *Environment and Planning B: Urban Analytics and City Science*. <https://doi.org/10.1177/2399808319828730>
- Jenkins A, Croitoru A, Crooks AT, Stefanidis A (2016) Crowdsourcing a collective sense of place. *PLoS ONE* 11(4):e152932
- Jiang J, Li Q, Tu W, Shaw SL, Yue Y (2019) A simple and direct method to analyse the influences of sampling fractions on modelling intra-city human mobility. *Int J Geogr Inf Sci* 33(3):618–644
- Jiang W, Wang Y, Tsou MH, Fu X (2015) Using social media to detect outdoor air pollution and monitor air quality index (AQI): a geo-targeted spatiotemporal analysis framework with Sina Weibo (Chinese Twitter). *PLoS ONE* 10(10):e141185
- Kim HJ, Chae BK, Park SB (2017) Exploring public space through social media: an exploratory case study on the high line New York City. *Urban Design International* 23(2):69–85
- Koskinen I (2003) User-generated content in mobile multimedia: empirical evidence from user studies. In: *Proceedings ICME*, pp 645–648
- Krumm J, Davies N, Narayanaswami C (2008) User-generated content. *IEEE Pervasive Comput* 7:10–11
- Lenders V, Koukoudidis E, Zhang P, Martonosi M (2008) Location-based trust for mobile user-generated content: applications, challenges and implementations. In: *Proceedings of the 9th workshop on mobile computing systems and applications*, pp 60–64

- Levy JM (2016) Contemporary urban planning. Taylor and Francis
- Li Q (2017) From geomatics to urban informatics. *Geomatics Inf Sci Wuhan University* 42(1):1–6
- Li X, Tu W, Shen S, Yang Y, Luo NX, Li QQ. (2017) Revealing spatial variation and correlation of urban travels from big trajectory data. *Int Arch Photogrammetry Remote Sens Spat Inf Sci* 42:53–57
- Li X, Zhang C, Li W, Ricard R, Meng Q, Zhang W (2015) Assessing street-level urban greenery using Google Street View and a modified green view index. *Urban Forest Urban Greening* 14(3):675–685
- Longley PA, Adnan M (2016) Geo-temporal Twitter demographics. *Int J Geogr Inf Sci* 30:369–389
- Martí P, Serrano-Estrada L, Nolasco-Cirugeda A (2017) Using locative social media and urban cartographies to identify and locate successful urban plazas. *Cities* 64:66–78
- Martí P, Serrano-Estrada L, Nolasco-Cirugeda A (2019) Social media data: challenges, opportunities and limitations in urban studies. *Comput Environ Urban Syst* 74:161–174
- Oliveira ACM, Botega LC, Saran JF, Silva JN, Neris VPA et al (2017) Crowdsourcing, data and information fusion and situation awareness for emergency management of forest fires: the project DF100Fogo (FDWithoutFire). *Computers, Environment, and Urban Systems* 77. <https://doi.org/10.1016/j.compenvurbsys.2017.08.006>
- Overeem A, R Robinson JC, Leijnse H, Steeneveld GJ, Horn BK, Uijlenhoet R (2013) Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophys Res Lett* 40(15):4081–4085
- Papadopoulos S, Zigkolis C, Kompatsiaris Y, Vakali A (2010) Cluster-based landmark and event detection for tagged photo collections. *IEEE Multimedia* 18(1):52–63
- Planet (2019) Data volume of OSM. <https://wiki.openstreetmap.org/wiki/Planet.osm>
- Santani D, Njuguna J, Bills T, Bryant AW, Bryant R, Ledgard J, Gatica-Perez D (2015) Commu-nisense: Crowdsourcing road hazards in Nairobi. In: *Proceedings of the 17th international conference on human-computer interaction with mobile devices and services*, pp 445–456
- Schnebele E, Oxendine C, Cervone G, Ferreira CM, Waters N (2015) Using non-authoritative sources during emergencies in urban areas. *Computational approaches for urban environments*, Springer, Cham, pp 337–361
- See L, Fonte CC, Antoniou V, Minghini M (2019) Volunteered geographic information: looking towards the next 10 years. *J Geogr Syst* 21:1–3
- Shelton T, Poorthuis A, Zook M (2015) Social media and the city: Rethinking urban socio-spatial inequality using user-generated geographic information. *Landscape Urban Plan* 142:198–211
- Tekalp AM (2015) Digital video processing. Prentice Hall Press, Upper Saddle River, NJ
- Thakuria P, Tilahun NY, Zellner M (2016) Introduction. *Seeing cities through big data: research, methods and applications in urban informatics*. Springer, Cham, pp 1–9
- Thomee B, Shamma DA, Friedland G, Elizalde B, Ni K, Poland D, Borth D, Li LJ (2016) YFCC100M: The new data in multimedia research. *Commun ACM* 59(2):64–73
- Tu W, Cao J, Yue Y, Shaw SL, Zhou M, Wang Z, Chang X, Xu Y, Li Q (2017) Coupling mobile phone and social media data: a new approach to understanding urban functions and diurnal patterns. *Int J Geogr Inf Sci* 31:2331–2358
- Tu W, Hu Z, Li L, Cao J, Jiang J, Li Q, Li Q (2018a) Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sens* 10(1):141
- Tu W, Cao R, Yue Y, Zhou B, Li Q, Li Q (2018b) Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *J Transp Geogr* 69:45–57
- Tu W, Santi P, Zhao T, He X, Li Q, Dong L, Wallington TJ, Ratti C (2019) Acceptability, energy consumption, and costs of electric vehicle for ridehailing. *Appl Energy* 250:147–160
- Valdez AM, Cook M, Langendahl PA, Roby H, Potter S (2018) Prototyping sustainable mobility practices: user-generated data in the smart city. *Technol Anal Strateg Manage* 30(2):144–157
- Venerandi A, Quattrone G, Capra L, Quercia D, Saez-Trumper D (2015) Measuring urban deprivation from user generated content. In: *Proceedings of the 18th ACM conference on computer supported cooperative work and social computing*, pp 254–264

- Wang D, Amin MT, Li S, Abdelzaher T, Kaplan L, Gu S, Pan C, Liu H, Aggarwal CC, Ganti R, Wang X (2014) Using humans as sensors: an estimation-theoretic perspective. In: Proceedings of the 13th international symposium on information processing in sensor networks, Berlin, Germany, pp 35–46
- Wang X, Zheng X, Zhang Q, Wang T, Shen D (2016) Crowdsourcing in ITS: The state of the work and the networking. *IEEE Trans Intell Transp Syst* 17(6):1596–1605
- Xu Y, Chen D, Zhang X, Tu W, Chen Y, Shen Y, Ratti C (2019) Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system. *Comput Environ Urban Syst* 74:184–203
- Xu Y, Chen D, Zhang X, Mei L, Luo X, Wei X, Hu C (2016) Crowdsourcing based description of urban emergency events using social media Big Data. *IEEE Trans Cloud Comput*. <https://doi.org/10.1016/j.compenvurbsys.2019.02.002>
- Yang D, Zhang D, Qu B (2015) Participatory cultural mapping based on collective behavior data in location based social networks. *ACM Trans Intell Syst Technol (TIST)* 7(3):1–23
- Yang D, Qu B, Yang J, Cudre-Mauroux P (2019) Revisiting user mobility and social relationships in LBSNs: a hypergraph embedding approach. In: Proceedings of The Web Conference (WWW'19), San Francisco, USA
- Yang J, Leskovec J (2011) Temporal variation in online media. In: ACM International Conference on Web Search and Data Mining (WSDM '11)
- Zhang F, Zhou B, Liu L, Liu Y, Fung HH, Lin H, Ratti C (2018) Measuring human perceptions of a large-scale urban region using machine learning. *Landscape Urban Plan* 180:148–160
- Zheng X, Chen W, Wang P, Shen D, Chen S, Wang X, Zhang Q, Yang L (2016) Big data for social transportation. *IEEE Trans Intell Transp Syst* 17(3):620–630
- Zhong C, Zeng S, Tu W, Yoshida M (2018) Profiling the spatial structure of London: From individual tweets to aggregated functional zones. *ISPRS Int J Geo-Inf* 7(10):386
- Zhou B, Lapedriza A, Khosla A, Oliva A, Torralba A (2018) Places: a 10 million image database for scene recognition. *IEEE Trans Pattern Anal Mach Intell* 40:1452–1464
- Zhu Y, Liu S, Newsam S (2017) Large-scale mapping of human activity using geo-tagged videos. In: Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, p 68

Wei Tu is an associate professor of the department of Urban Informatics, Shenzhen University. He received his Ph.D. degree in photogrammetry and remote sensing from Wuhan University. He has been a visiting scholar at the Senseable City Laboratory, MIT. His research interests include urban informatics, urban sensing, and trajectory analytics.

Qingquan Li is a Professor at Shenzhen University. He is the President of Shenzhen University, Shenzhen, China, and the Director of the Guangdong Key Laboratory of Urban Informatics. He is a member of the International Eurasian Academy of Sciences. He has received several national awards, including the State Scientific and Technological Progress Award, the state Technological Invention Award, and the Ho Leung Ho Lee Award.

Yatao Zhang received his B.S. degree from Sun Yat-sen University. He is currently working toward an M.S. degree in GIS in the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China. His research interests include multi-source spatiotemporal data fusion and urban sensing.

Dr. Yang Yue is a Professor of Urban Informatics, with an interdisciplinary background in Geomatics, GIS, and urban studies, and especially focuses on urban big data studies. She has published over 50 peer-reviewed research papers and serves as a committee member in local and international GIS, computer, transportation, and urban planning academic advising organizations.

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.