Open Data, Crowdsourcing, and City Planning

A Novel Perspective on Public Participation in Planning and Public Governance

Mingrui Mao[1*] Ying Long[1]
[1]Planning information centre Beijing Municipal Institute of City Planning & Design
Beijing, China
maomingrui@gmail.com

***Abstract*—With the development of information technologies, open data released by central and local governments and crowdsourcing means are emerging in China. In this article, we first address this phenomenon by describing open data source, open data, and crowdsourcing datasets. We secondly address case studies on using open data are elaborated, most of which are developed by a recently-established network city lab Beijing City Lab (BCL, http://longy.jimdo.com). Lastly, we address the potential opportunities for public participation in urban planning and public governance for urban management in detail.**

Keywords: open data; crowdsourcing; public paticipation; city planning; public governance; Beijing City Lab (BCL)

I. OPEN SOURCE, OPEN DATA AND CROWDSOURCING

Open source and crowdsourcing are two popular terms which become popular with Internet evolvement. The Open Source indicates the open source codes and describes a method for providing final source materials in product output and development. Before a computer is invented, the sharing of technology information has existed for a long period of time. An example of which is the sharing of recipes occuring at the beginning of human being culture. The open source quickly develops with rising and prosperity of Internet and incurs a series of "open" activities such as open data.

The open data indicates some elected and permitted data without restriction from copyright, patent right, and other management mechanisms. This data can be open to the social public and anyone can freely publish and use them regardless of publishing or use for other purposes. The target of open data promotion is similar to it of other "open" promotions, such as open source, open contents, and open knowledge. The kernel idea behind Open data has been formed permanently, but the term "Open data" only emerges in the modern time and is well-known due to emerging of Internet, especially setup of Open data governmental organizations such as Data.gov.

Crowdsourcing (mass outsourcing) is a new production organization form caused by Internet, e.g. OpenStreetMap (OSM) [1]. Although this organization form has emerged for a long period, the term "Crowdsourcing" is very young, which emerged in 2006. It describes a business pattern, namely an enterprise assigns its work via Internet to find originalities or solve technical problems. Under control of Internet, these organizations can utilize originalities and capabilities of volunteers, who grasp sills to complete tasks, are willing to work in their spare time, are satisfactory for small rewards for their service, or no reward for a temporary period, and expect to get more rewards in future. Especially for the software industry and service industry, it provides a brand-new pattern for labor organization.

The open source and openness promotions drive huge changes in many areas such as IT, education, and government governance. More and more platforms open self API for other applications to access self-data. More and more open classes emerge like mushrooms after rain. More and more libraries open self-resources and more and more governments join in the list of the open governments.

Crowdsourcing is not only one business pattern, but it also effectively changes public living. In a general meaning, public comment websites, such as Wikipedia and Sina microblog, are well-known crowdsourcing projects. For the recently focused MH370 search, DigitalGlobe America Satellite Corporation launches mass to determine high-resolution image data in thousands of kilometers via the crowdsourcing pattern in order to speed up search. Until March 14, about 3 million volunteers mark 2.9 million feature points and view satellite image data 250 million times on the Tomnod CrowdSourcing platform (http://www.tomnod.com/nod/challenge/malaysiaairsar2014?source=malaysia). Each pixel of the image is viewed by nude eyes at least 30 times. The crowdsourcing activity for MH370 has successfully become the largest crowdsourcing activity in the world.

II. URBAN STUDIES USING OPEN AND CROWDSOURCING DATA

Some city researchers with geography and city planning background study the city from different scales and views based on the open geography data, social media data, and check-in data, including macro research such as city form, area association, and micro research, such as individual behavior pattern.

Dr Ying Long from Beijing Municipal Institute of City Planning & Design launched a network city lab Beijing City Lab, BCL, http://longy.jimdo.com/ at the end of 2013. BCL is dedicated to quantification of the city development dynamics with cross-subject method and research on city science. BCL is the first open quantitative city research network in China, which invites scholars to expound the latest opinions on city research in manners such as publish working paper and provides open quantitative city research data for the science

research groups via data sharing. BCL is a research platform which complies with openness, sharing and crowdsourcing spirit, which is recognized by multiple experts and scholars in China and abroad (refer to BCL honorable members for details). BCL converges a large number of city research achievements based on open data and crowdsourcing data. The following part lists several typical achievements:

OpenStreetMap (OSM) is a map drawn and edited by users according to handheld GPS device, aerial photographs, satellite images, event understandings on related areas. Ying Long and Xingjian Liu (BCL working paper 16) used road networks in OSM and ubiquitous points-of-interest to partition geographical space, delineated urban parcels and inferred urban function, development density and land use mix for each urban parcel for 297 Chinese cities, which provides foundational data for quantitative urban studies for Chinese cities [2]. Based on it, BCL also carried out extended research on quantitative urban studies for all Chinese cities, and proposed a new urban study solution "mega model". For details, refer to the following part.

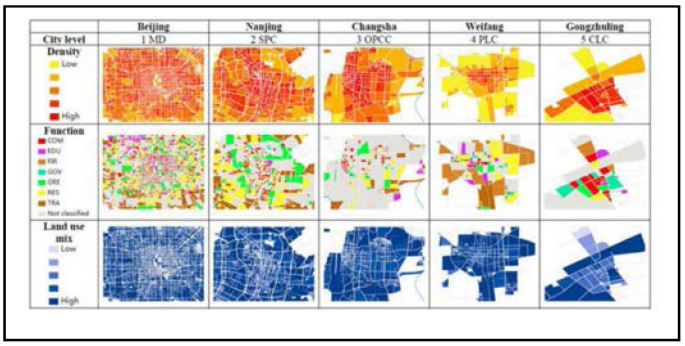

Figure 1. Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest. [2]

Based on 15 million Sina Microblog check-in records during one year in Shanghai, Wu lun, Zhi ye, Sui Zhengwei and Liu yu combine the activity-based analysis with movement-trajectory approach to model the intra-urban human mobility [3]. Long Ying analyzes the mixing degree of Beijing city by using check-in data and POI data from Sina Microblog [4]. Mao Mingrui analyzes the connection network and space distribution of the planners of China on Sina Microblog [5].

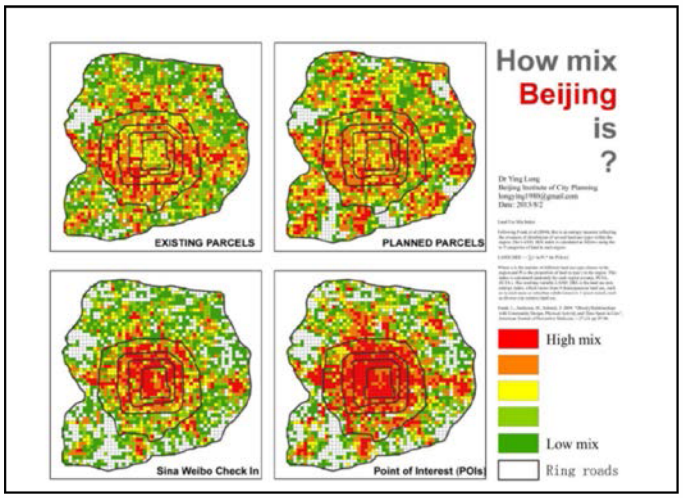

Figure 2. How mix is Beijing? [4]

Based on POI data at the websites such as Dianping.com, Liu Xingjian visualizes distribution of gastronomy in China [6]. Wang Peng analyzes relevance of the creative industry and coffee house in space distribution (http://weibo.com/1425406277/A3XqBaAQh).

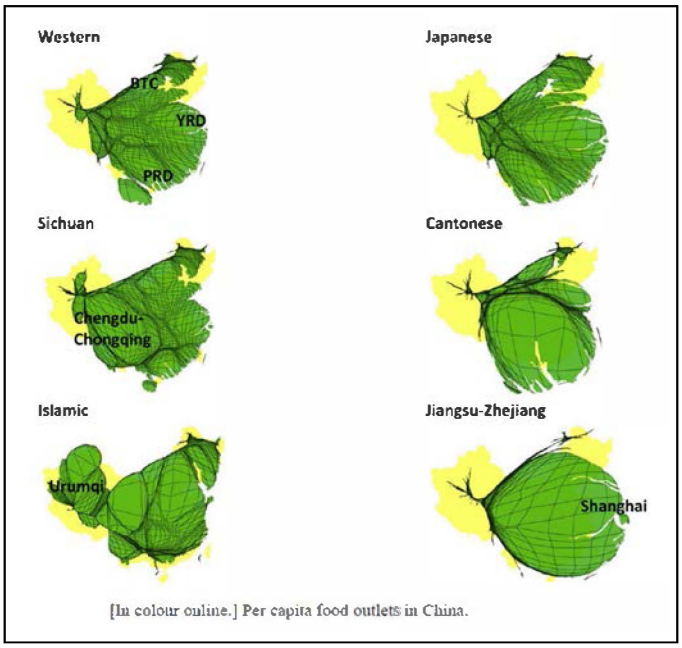

Figure 3. Visualizing urban gastronomy in China. [6]

Considering huge role of open and crowdsourcing data on the city planning and quantitative research on cities, Mao Mingrui, Wu Yunchao and Long Ying study the intelligent crawling of multi-source social data and plan application, arrange sources, acquisition methods, handling means, plan application ideas of the valuable open data and crowdsourcing data for city planning and research on the Internet, and form systematic research achievement (for details, please refer to BCL Slides 5) [7].

TABLE I. WEBSITES OF OPEN DATA AND CROWDSOURSING DATA

Type	Name	Website
OD: Country	Open source Holland	Oscity.nl
OD: City	Open source city plan	github.com/
OD: City	List of open source cities	futureeverything.org/2012/11/open-data-cities/
OD: business website	Data class	www.datatang.com
OD: government	Beijing government office data resource website	www.bjdata.gov.cn
OD: government	National data	data.stats.gov.cn
CS: VGI	Open Street Map	www.opentreetmap.org
CS: VGI	Global cities Data	Download.bbbike.org/osm/bbbike
CS: snapping	sightsmap	www.sightsmap.com

address		

Remark: OD: Open Data, CS: crowdsourcing

III. INFORMATION IS POWER: NEW VIEW FOR PUBLIC PARTICIPATION IN PLANNING AND PUBLIC GOVERNANCE

The application meaning of the open data and crowdsourcing data in city planning and city research is far more than its learning value or technology innovation in a degree, and is deeply affected by city planning and government governance in China.

The compiling of traditional city plans highly depends on foundational information from several official channels, such as survey map from surveying and mapping department, statistical data of statistics bureau, government office data of different commissions, offices and bureaus, and small-scope sampling survey data. Generally, this data can only display very macro, simple, and rough statistics information, lose many details and basically show objects instead of persons due to restricted statistics diameter of governments, data barriers between departments and rough degree of government management. Although a plan compiling pattern is summarized from multi-year plan practices based on these data in China, it does not indicate that the plan and decision based on these data and patterns are scientific.

At the same time, the public participation threshold of city planning is very high due to different reasons. Now most of limited public participation means such as plan publicity required by the laws and regulations are formalities, so it is difficult for the public to participate in city planning and play actual influence on plan decision.

The above realities can be divided into two information monopolization. One is "supply" monopolization from the governmental departments. The plan is only compiled based on the limited foundational data monopolized by the governments, so it objectively restricts science of plan decision. Another is "requirement" monopolized by the governmental departments. Only the governmental requirements can be heard in plan compiling, decision and evaluation and it is difficult to observe the public requirement feedbacks due to the insufficient public participation methods and degree. Monopolization of two information leads to full control of the public power on the city planning and makes the city planning become one dependent technical tool of the government power.

Emerging open and crowdsourcing data and generations of big data processing and mining methods brings a new dawn for changing the above layout.

On the one hand, the open government trend has started to affect China. Since "information publication regulations of PRC" was issued in 2007, the publication degree of governmental affairs information keeps improving. Since 2010, several governmental data websites are online. E.g. the "national data" (http://data.stats.gov.cn) of National Statistical Bureau and Beijing governmental affairs data resource network (http://www.bjdata.gov.cn) are excellent representatives. The open knowledge Foundation recently issues census results of open governmental data in 2013. The rank of China's government is not low, and it is superior to several countries in Europe. Meanwhile, more and more social websites open self-data or interface. Information openness makes governmental behaviors more transparent to citizens, so the public gets the "knowing right". The governmental affairs information opened by the governments also breaks through the monopoly of governmental information, and openness of the social data breaks through governmental monopoly on the information publishing channels. The social media break through controls the news, media, and social opinion. The air PM2.5 index, tested by the civilians and third-party organization, forces the government to publicize corresponding public information. The city planning and research needs to not fully depend on traditional data acquisition means. Some research on the above-mentioned BCL platform publishing embodies huge value of open data on city plan and research.

On the other hand, the crowdsourcing brings a new public participation form, and it may drive a change of the plan decision and government governance. The main behavior body of crowdsourcing is people, and the crowdsourcing data embodies activity information on persons. Dianping.com opens data interface so the public and city researchers not only completely know information on different utility service facilities in cities in real time, but also knows facility use frequency and evaluation of public. Public check-in data enables city researchers to observe space and time distribution of person flow and person behavior pattern in different areas of the city. The shooting position information uploaded by the public enables city planners and researchers to know city image of different populations and time and space change of the hot tourism areas. The social media, such as microblogs, provides a platform to sense and monitor public opinions and operation dynamics in cities and event emotion of citizens. From the view of general sense, the behaviors of citizens, such as bus taking, smart card data for subway use, and mobile phone use, provide a platform for researchers and managers to sense citizen activities. If this sense is regarded as a project with a purpose, the above behaviors of the citizens can be crowdsourcing behaviors. Baidu Migration, which was put into operation on 2014 Spring Festive, can show the visual results based on the mobile phone position behaviors of the public.

If the open data breaks through the information monopoly, the crowdsourcing creates data foundation for sensing social and person activities, which cannot be acquired in traditional plan preparation and plan research. Since the city's plan proposes the "human-oriented" slogan, the plan first truly meets the data conditions for humanism plan practice after several years. The public participation of the city plan also changes from the past small-scope questionnaire survey and plan publication from up to down dominated by the planners and city governors to spontaneous, deserved and popular behaviors from down to up. Crowdsourcing and valid analysis on crowdsourcing data reduce the participation threshold of the citizens in the public affairs.

Although some disputes exist, E.g. the crowdsourcing is a voluntary behavior and the crowdsourcing participants can not represent all citizens. Planned research and decisions based on the crowdsourcing data may deviate, but we think that openness, sharing and crowdsourcing spirit are matched with the modern citizen sprit. The crowdsourcing participants can be sensed and monitored much by city planners, researchers, and

government deciders, so the public policies can be made based on the shared data and further generate decisions favorable to the sharers, which is rewards and encourages for the public to participate in public affairs and embodies the social humanism, pluralism, openness, openness and participation of the citizens.

Therefore, we only face a problem. When the open and crowdsourcing big data knocks on the door, how will the planners and city researchers embrace this "big" happiness? When the power flows from the government to the society and public with information, how will the planners and government governors change the pattern and governance methods?

Recently, Michael Batty published a book named as the quantitative research on the city named "The New Science of Cities", which points out that several sciences on city exist. The "new sciences" is named because the technologies and tools used by this science are relatively new.

At the same time, the second author of this paper Dr Ying Long proposes "mega model" concept by analyzing the city model development trends of the big data era, which is a quantitative research tool for cities and units, and is driven by large-scale data, mainly utilizes simple, and direct modeling method, consider large-scale and fined simulation unit, and represents a new research paradigm (refer to BCL Slides8 for details) [8]. The traditional model studies compromise between the scale and simulation units (restricted by data and computing capability), but the large model considers the research scale and simulation unit (large geographical extent and fine granularity). The general standard of the mega model is that the space scale is a city group or larger (E.g. whole China). The simulation unit depends on the space scale such as land mass, block, or street (the simulation area includes over 10000 research units). The modeling of large models is mainly based on the traditional modeling idea and is more simple and intuitive. On the whole, a mega model studies city and area problems from the micro view and is dedicated to solution of scientific problems, so it is a science problem. The mega model reflects the individual activities and movement and can alleviate the gap between technologies and digit in the medium-size and small-size cities (most past city models are established for large cities with advanced data and technology). Now BCL has carried out some research on the mega models for all large-size, medium-size, and small-size cities in China such as nationwide expansion model of the land block scale, city constructed area identification, area block boundary and development type and strength re-modeling, inter-city traffic network analysis and simulation model in China, time and space evolution analysis models of city and town patterns in China, city population data synthesis, and resident living quality evaluation in China based on the area block scale, air pollution exposure evaluation, city area scope division of main cities, and city group growth evaluation in progress.

Regardless of the "new city science" [9] or the "mega model", they point at the answer to the above problem. One main background of the reform of city planning and government governance is openness and crowdsourcing. One main task for reform planning is to carry out city research and plan innovation with new city science, realize fine simulation of cities with a mega model and further realize the fine governance of cities.

The director Lei Zhang from digital governance and mobile governmental affairs of Fudan University thinks it as a "crowdsourcing" pattern. In the crowdsourcing pattern, the citizens become the 'cooperator' of the government. The citizens cooperate with the governments to 'provide' public service and not just 'consume' services independently provided by the governments.

Finally, it comes to the word "Information is power".

ACKNOWLEDGMENT

The authors would like to express appreciations to all BCL members for their helps.

REFERENCES

[1] M. Haklay, P. Weber, "OpenStreetMap: User-generated street maps," Pervasive Computing, 2008, vol. 7, pp. 12-18.

[2] Y. Long, X. Liu, "Automated identification and characterization of parcels (AICP) with OpenStreetMap and Points of Interest," arXiv, 2013, preprint arXiv:1311.6165.

[3] L. Wu, Y. Zhi, Z. Sui, & Y. Liu, "Intra-Urban Human Mobility and Activity Transition: Evidence from Social Media Check-In Data". PloS one, 2014, 9(5): e97010.

[4] Y. Long, X. Liu, "How mixed is Beijing, China? A visual exploration of mixed land use," Environment and Planning A, 2013, vol. 45, pp. 2797–2798.

[5] M. Mao, "Analysis on social relationship of CAUPD, THUPDI and BMICPD on Sina Microblog", 2013, vol. 5, pp. 178-180.

[6] X. Liu, "Visualizing urban gastronomy in China," Environment and Planning A, 2014, vol. 46, pp. 1012-1013.

[7] M. Mao, Y. Wu, Y. Long, "Intelligent data crawling and planning applications in Beijing," Beijing City Lab, Slides 5, 2014.

[8] Y. Long, "Mega model," Beijing City Lab, Slides 8, 2014.

[9] M. Batty, The New Science of Cities. MIT Press, 2013.