

为城市体检 ——大数据在城乡规划中的应用

PHYSICAL EXAMS FOR THE CITY — APPLICATIONS FOR BIG DATA IN URBAN AND RURAL PLANNING



王鹏
北京清华同衡规划设计研究院有限公司技术创新中心副总工程师、高级规划师

Peng WANG
Associate Chief Engineer and Senior Urban Planner at Technology Innovation Center, Beijing Tsinghua Tongheng Urban Planning and Design Institute

采访日期 INTERVIEWED DATE: 2015-06-11
中图分类号 / TU986.2 文献标识码 / C

摘要

本次访谈的对象——高级规划师王鹏，更倾向于对大数据采用较为宽泛的定义，将其视为一种方法论。结合具体实践案例，他就目前大数据对城市规划师的影响、从业人员对大数据的误解、大数据面临的困难，以及大数据的未来发展进行了深入分析与解读。

关键词

大数据；城市规划；体检；方法论

ABSTRACT

In this article, the interviewee, senior urban planner Peng Wang, suggests a broader definition for Big Data, considering which as a methodology. Combining introduction of case studies, he offers insights on the issues of how Big Data influences planners' practice, the misunderstandings in planning industry about Big Data, the obstacles in the research and application of Big Data, as well as Big Data's future development.

KEY WORDS

Big Data; Urban Planning; Physical Exam; Methodology

采访、编辑 涂先明 田乐

译 吴露 萨拉·雅各布斯 田乐

INTERVIEWED AND EDITED BY Xianming TU Tina TIAN

TRANSLATED BY Xiao WU Sara JACOBS Tina TIAN

在您看来“大数据”最核心的特征是什么？

王鹏（以下简称“王”）：大数据的核心特征之一是“海量”，也就是说数据量需要达到一定级别。但这一点也需要辩证地来看，因为每个时代、每个行业都有着对“海量”的不同界定。计算机领域一般认为需要至少T级以上的数据才可以称得上大数据——通俗来讲就是无法用Excel来进行处理的数据。如果按照这样的界定，城市规划行业中目前基本不会遇到这样级别的数据。所以，我个人更倾向于对大数据采用较为宽泛的定义，将其视为一种方法论。相较于传统的抽样调查方法——试图用小样本量去还原全貌，实际会存在很大的误差——而言，大数据则是一种接近全样本的调查方式，使得我们可以更加准确地观察到全貌。

大数据可以大致分为两种类型。一类是来自物联网、互联网的数据，这些数据的数据量非常庞大且源源不断地生成，这类数据毫无疑问是大数据。还有一类是对于传统数据的再挖掘。举例来说，过去我们只能对某一个县的情况进行纵向分析，或者对少数几个城市进行横向对比；现在我们可以建立全国所有县的各项统计数据的数据库，从而能够对所有县的情况进行对比，寻找其中的规律。虽然这种横向对比的思想在全国城镇体系规划中已然存在，但是大数据可以使这样的研究更加便捷、更具有可操作性。

其次，大数据也会为我们的工作方式带来变革，例如，我们研发的移动端调研工具，可以在调研前将场地中关于建筑、道路

等的基础信息录入到平台中，一个项目组十几位成员带着安装有这个工具的手机到达现场后，每个人用手机所拍摄的照片或视频，都可以集中到这个平台上，并且我们还可以对拍摄的地点等信息进行分析。针对一个项目，我们可能要拍摄几千张甚至上万张照片，过去我们只能从中挑选出部分作为代表，而现在我们可以系统地归类所有图像信息。此外，在调研时这些信息可以自动录入，在调研结束后直接同步到系统就可以了。

此外，大数据也将带来更直接、更广泛的公众参与。例如我们可以开发一些APP，让公众来了解城市中的景观、建筑、文化遗产，并可以对其进行评价。这样有助于让公众从过去的被动参与变为主动参与。

从目前的实践层面来看，是否可以说大数据的方法更适用于大尺度的规划？

王：大数据中的很大一部分都是有关人的行为的数据，这为我们提供了一个观察人的重要视角。过去，我们往往更关注城市中的建筑与空间，而忽略了空间中的人。这是由于过去种种技术局限，我们无法对在城市中的位置进行准确定位。但是现在无论是借助社交网络，还是移动终端，都可以帮助我们去了解每个人在每个时间点位于空间中的位置与行为，大数据使我们看到了采用传统的调查方式所无法观察到的内容。尤其重要的一点是，我们的城市规划已进入存量规划的阶段，大数据能有效地帮助我们发现建成环境的问题，我们可以借助其进行很多的探索。例如，我们可以在城市尺度下，分

析城市职住位置关系、城市的空间结构特征等，甚至可以在全国尺度，来分析人群在城市间的移动，从而辨析城市之间的关联性。

目前来说，由于我们所能获得的很多数据的精度还不够支撑小尺度的项目，例如定位数据中我们能获得的手机信令数据只能精确到二三百米的范围内。但这并不意味着小尺度的项目与大数据“无缘”，例如，大数据可以为我们提供一种设计的反馈回路：我们可以针对一些已建成的景观设计一些APP，来观察使用者在这些景观中的行为轨迹以及对景观的评价，从而能够对项目进行进一步的改善。

从现有的研究成果来看，有没有哪些成果完全推翻了过去的认知？

王：这样的案例有很多，最典型的案例是交通的出行调查。过去我们对城市居民就业与居住的位置进行调查，采用的是抽样调查的方式，并试图恢复所有居民在城市中移动的状况，进而估算城市的交通需求。传统的抽样调查方法在理论上是可行的，但过去缺乏对其结果校验的途径。但是，现在城市中几乎“人手一机”，保守地说手机信令能够覆盖城市中百分之七八十的人群，手机信令可以提供一个接近于全样本的数据。从现有的一些分析结果来看，传统的交通调查结果远不如手机信令的分析结果准确。

另一个案例是我们正在进行的张家口草原天路项目。这个景区刚刚兴起，在网络上由于游客口口相传，导致去年游客爆棚，很多问题也随之而来，例如在十一黄金周期间

PHYSICAL EXAMS FOR THE CITY — APPLICATIONS FOR BIG DATA IN URBAN AND RURAL PLANNING

就出现了严重的交通拥堵。因此，管理者委托我们对景区进行重新规划，要求我们在今年十一黄金周到来之前就提供解决方案。但是这个景区根本就没有游客量的历史数据，我们只知道出现了堵车、服务设施不足等问题，但是这些问题是什么时段、在哪些地点出现的，我们不得而知。由于在非黄金周期间景区的游客量很低，实地调研也无法对我们提供有效的指导。如果用传统的旅游规划方式，没有历史数据就难以进行分析了。但是大数据可以帮助我们去了解当时的情况。例如，我们可以借助去年十一黄金周期间的微博，来分析游客对哪些景点感兴趣、会在哪些地点聚集，或者根据游客发送微博的地点来推测拥堵的地点。虽然说并不是所有的游客都会使用微博，但其至少为我们提供了一些线索。但是微博的用户目前也在逐渐衰减，来自微博的样本量远不达到我们的预期。所以未来我们还是要尽力去搭建更具针对性的公众参与平台。此外，我们还要探索多种获取数据的途径，例如在这个项目中，我们借助了某互联网公司的LBS定位数据——互联网公司的后台网络可以对绝大多数的智能手机用户的位置进行记录——这份数据能够提供很明确的游客路线、游客停留的地点等信息。清楚了问题所在可以说就已经完成了解决问题的大部分工作了。

您认为规划行业从业者对于大数据存在着哪些认知误区？

王：对于绝大多数的行业从业者来说，对大数据的概念还是非常模糊的。一个很重

要的原因是目前大家在工作中难以接触到大数据，因此容易过于“神化”大数据。大数据确实为我们提供了思考行业问题的新视角，但并不意味着大数据的应用会带来一场规划界的革命。我常常用医学的诊疗手段来类比：我们有了CT，有了X光，但是我们依旧需要医生依据其专业经验、结合病人的临床表现来决定具体的治疗方案。在存量规划阶段，我们更多的工作并不是“做手术”，而是“做体检”，大数据正是对城市的一种“体检”，有助于我们识别城市中存在的问题，但我们依旧需要规划师通过制定具体方案来解决这些问题。

另外一个误区是，很多人在了解大数据目前的成果之后，认为大数据不过如此，并会对大数据的未来发展产生过于武断的否定结论。这其中的一个原因是由于只有少量的人在从事这方面的研究，拥有地理学、城市规划、景观设计学背景，能够对城市问题作出专业判断，同时擅长计算机技术的人才非常难得。在目前这个实验、摸索阶段，肯定有很多不成熟之处。但我相信未来一定会有越来越多的大数据成果会更新我们对于城市问题的认识。

目前来说，大数据的研究与应用所面临的困难有哪些？

王：在传统的增量规划中，我们会根据已经发生的事实所积累的数据，寻找其中的规律，然后建立一定的模型，来预测未来的发展趋势——其中非常典型的一个例子是我们对于城市人口规模的预测。而对于大数据

来说，其现阶段被更多地应用于反映历史和现状，而非预测未来的发展趋势——目前来说做到这一点还是非常困难的，因为，正如我前面所说，大数据中的很大一部分都是有关人的行为的数据，要做到对于人的行为进行精确记录、发现其中的规律，并分析怎么样的设计干预会带来怎样的使用结果是需要长期的数据积累和大量的精细化研究的，而我们目前并不具备相关的技术与资源支持。

在未来，规划师是否只有玩得转数据，才能玩得转规划？

王：借用现在很流行的一种说法：“未来的规划分两种：一种是下社区，一种是玩数据。”下社区是指近距离地观察人的行为，数据则是更科学地分析人的行为，所以无论是下到社区去观察，还是用数据来观察，都意味着未来的城市规划将越来越多地关注人的行为。我们将更重视物质空间“好不好用”的问题，而非“好不好看”。LAF

1. 北京市文化遗产保护应用程序。该APP对北京市文化遗产的重要性等级进行了标签化处理：红色的为国家级文化遗产，蓝色的为市级文化遗产，绿色的为区级文化遗产。
2. 北京市文物保护点应用程序。该APP建立了文化遗产索引与评分体系。

1. Application for cultural heritage preservation of Beijing. This App labels all the monuments and sites in terms of the hierarchy of importance: national heritages (in red), municipal heritages (in blue), and district heritages (in green).
2. Application for cultural heritage preservation of Beijing. This App provides index and ranting services that allow users reviewing all the historic monuments and sites in Beijing.

What are the characteristics of Big Data?

Peng WANG (WANG hereafter): One of the core characteristics of Big Data is scale, which means that the quantity of data needs to reach a certain level. This also requires a certain dialectical approach to what is considered “big”, because every era, and every industry, has a relative definition of “big”. In computational technology only data at a T level can be called Big Data, which, saying informally, is larger than what can be processed using Excel. At present, urban planning can hardly reach this level. I prefer a broader definition for Big Data, and one that concerns methodology. Big Data should be used to understand a full sample, so we can use the whole picture accurately rather than traditional sampling methods that draw from small sample sizes.

There are two types of Big Data: the Internet and the Internet of Things; and the re-mining of traditional data. Data generated from the Internet is huge and endless. The second draws larger samples sizes from traditional data methodologies. For example, rather than analysis of a single county, or lateral comparisons for a few cities, we can now compare statistical data and find patterns for all counties in the country. Lateral comparison already exists in the National Urban System Planning, but Big Data makes such research more

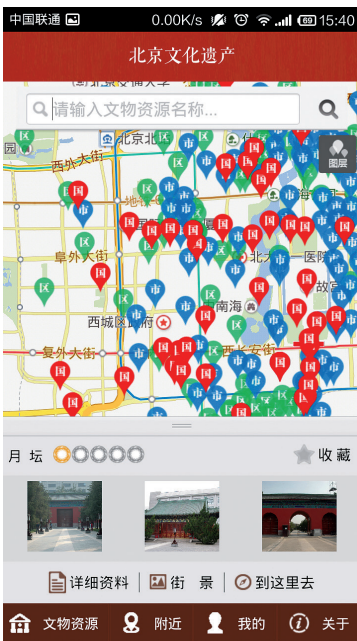
convenient, feasible and effective.

Meanwhile, Big Data is bringing changes to the way we work. For example, the mobile field research tool we have developed can record and aggregate basic information about the existing buildings and roads to a shared platform in advance. Photos and videos taken on site can be uploaded to the platform and can be searched or analyzed at a later time. We may have thousands of photos for one project, but in the past we could search and select only a small portion. Now we can systematically classify all image information. Additionally, information can be recorded automatically and directly synchronized to the system.

Moreover, Big Data will also bring more direct and extensive public participation. For example, we can develop online applications for the public to learn about the city’s landscape, architecture, and cultural heritage, with evaluation or comment function, encouraging them to participate in a more active way.

Are Big Data methods more suitable for large scale planning efforts?

WANG: Most Big Data is related to human behavior, which offers us an important perspective to observe human patterns and practices. In the past we have tended to be more concerned



with architecture and space in cities, rather than the people who live in them. We were unable to accurately locate people’s locations within the city. Social networking now helps us understand spatial position relative to behavior at any point in real-time. Big Data enables us to see things that cannot be observed using the traditional methods. Importantly, urban planning has entered an inventory planning stage meaning that Big Data can help us to identify the problems of our built environments. At the urban scale we can analyze spatial relationships between work and life, the spatial characteristics of cities, migration between cities, and inter-city relationships.

The data currently available to us is not accurate enough. Location data, for example, is only accurate at a range of two to three hundred meters. However, Big data can still support small-scale projects’ feedback loop, an application for user behavior tracing and evaluations can be developed for instance.

Drawing from current research, is there anything that has overturned previous research?

WANG: There are many cases. The most typical examples are transportation surveys. In the past, when we estimated urban traffic needs we investigated the commuting distance through traditional sampling methods, attempting to generate a city-wide

pattern of commuting and traffic shifts. Although traditional sampling methods were feasible in theory, but in the past there was a lack of verified results. Now mobile phone sampling can cover seventy to eighty percent of people in cities that offers a possibility of forming a nearly full sample by utilizing urban cell phone signaling. With the use of cell phone data the traditional traffic study outcomes become far less accurate than mobile phone signaling analysis.

Another example is our Zhangjiakou Prairie Heavenly Road project. Word of mouth caused a boom of visitors last year, with chaotic traffic problem in holidays and festivals or specific events. The resort commissioned us to create a new master plan before the National Day Holidays next year. However, the resort had no record of visitor patterns; questions such as what time and where congestion occurred or service facilities in short supply was unknown to us. In addition, the results from daily observation cannot substantially contribute to the master plan because of the huge difference of visitor numbers between holidays and typical days. It would be not feasible to conduct a traditional planning without analyzing historical data. Big Data helped us understand the situation. For example, we used last year’s Chinese National Day Holidays’ microblogging (Weibo) to understand which attractions tourists are most interested in and where they gather. Although not all visitors used microblogging, it was an easy way

to give us some clues. However, Weibo is currently on the decline and the number of microblogging samples was far below our expectations. In the future we will need to build a more targeted platform for public participation and to explore new approaches for data collection. For this project we were also using a Location Based Service (LBS), supported by an Internet company, that can record the behavior and location of a majority of smart phone users — identifying the problem is a big step to solve that.

What misunderstandings do you think the planning industry has about Big Data?

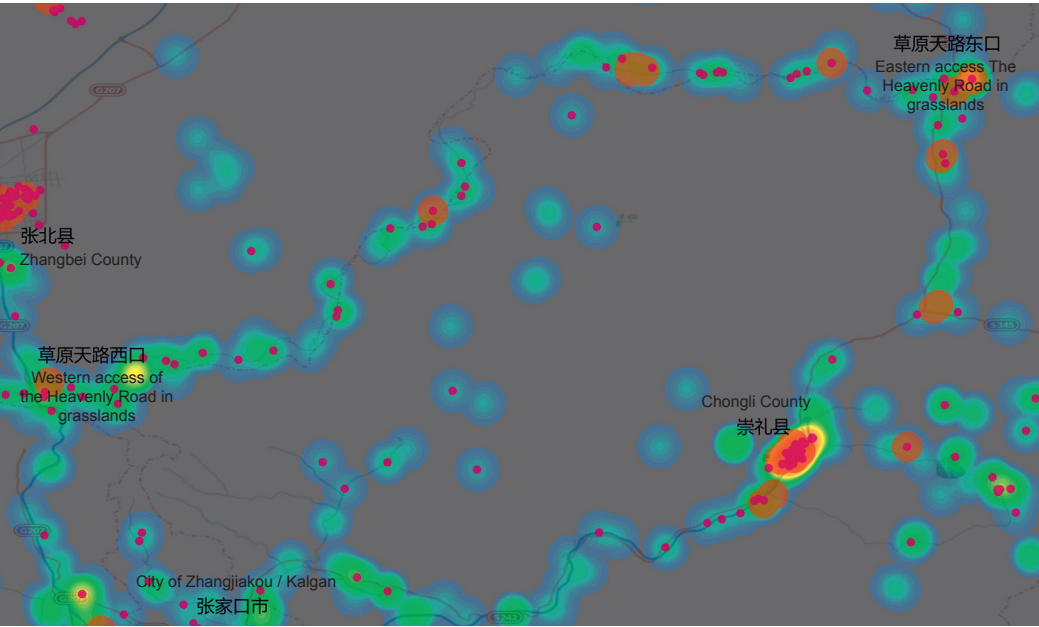
WANG: For many, the concept of Big Data is still very vague. It is largely due to that, at present, Big Data has not been directly involved in most practitioners’ work, and easily been deified by some planners who do not know very much about it. Big Data indeed provide us with a new perspective to think about industry issues, but that does not mean it will bring a complete revolution to the planning industry. I often use a medical analogy: despite all the medical tests and technologies, we still need a doctor to determine the treatment based on their professional experience. Big Data is the kind of “physical exams” for cities. It helps us identify the problems, but we still need planners to address the specifics. In the current inventory planning stage, the most important job to planners is to take “exams” for cities,

rather than to take “surgeries”.

At the same time, some planners devalue the existing efforts what Big Data can do now, and even with poor prospects of Big Data’s future development. One of the reasons is that, the number of experts proficient in Big Data is quite small; talents with geography, urban planning, and landscape architecture background, who are also good at computer technologies are very rare. In this experiment and exploration stage, there are certainly many things that are immature. But in the future there will continue to be more Big Data outcomes that will help expand our knowledge of urban issues.

What obstacles do you think in the research and application of Big Data?

WANG: In traditional incremental planning, data is gathered from facts to find regularities, build models, and predict trends. A very typical example is urban population size prediction. Currently, Big Data has usually been used for recording historical and current situation, but less in predicting trends; because, as I mentioned above that most Big Data is related to human behavior, and we have not had sufficient supports from current technologies and resources to conduct high-accuracy records of human behavior, find out specific patterns, and



generate feasible design outcomes and evaluate the usage. We need to conduct much more long-term, high-accuracy research in future.

In the future, will planners have to be good at data to be good at planning?

WANG: The future of planning will be twofold: community based and data based. Community outreach relies on close observations of human behavior, while data based research uses scientific analysis. So whether community or data based, the future urban planning will increasingly focuses on human behavior and how space is used, rather than how it looks. **LAF**

- 2014年10月1~7日草原天路附近微博发布地点分析图。由黄-蓝渐变区域为微博发布地点，数据显示，除了周边城镇（张北县、崇礼县、张家口市），草原天路沿线有一定数量的微博数据，并且集中在景点周边。图中橘色区域表示在微博内容中提及的发生交通拥堵的区域，主要集中在天路入口区域及部分路段。图中红色点则表示在微博中发布图片数量多于7张的地点：这些地点在草原天路沿线均匀分布，主要集中在景点附近。在一定程度上，发布的照片越多，说明游客对于风景的喜爱程度越高。
- Weibo sending map of the scenic area of the Prairie Heavenly Road, during October 1 to 7, 2014. The yellow-blue spots indicate the locations of Weibo sending. Besides the ones in the neighboring areas (Zhangbei County, Chongli County, and the City of Zhangjiakou), the sending spots were located around scenic places along the Heavenly Road. The tangerine spots indicate the places at where traffic congestions were reported, clustering in the access areas and some certain sections of the Road. The red spots indicate the locations at where the Weibos with more than seven photos were sent. Generally, these locations were evenly distributed along the Road, clustering in the scenic places. In a certain degree, the places where more photos visitors sent are more popular.