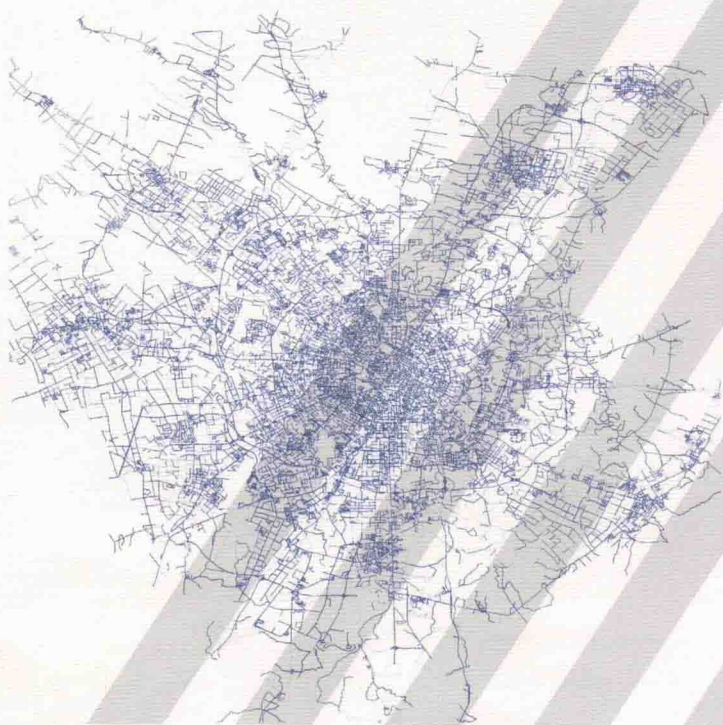




住房和城乡建设部土建类学科专业“十三五”规划教材
高校城乡规划专业规划推荐教材

城市规划大数据 理论与方法

龙瀛 毛其智 著



中国建筑工业出版社



住房和城乡建设部土建类学科专业“十三五”规划教材
高校城乡规划专业规划推荐教材

城市规划大数据 理论与方法

龙 瀛 毛其智 著

中国建筑工业出版社

图书在版编目(CIP)数据

城市规划大数据理论与方法 / 龙瀛, 毛其智著. —北京 :
中国建筑工业出版社, 2018.8
住房城乡建设部土建类学科专业“十三五”规划教材
高校城乡规划专业规划推荐教材
ISBN 978-7-112-22592-7

I . ①城… II . ①龙…②毛… III . ①互联网络—应用—
城市规划—高等学校—教材 IV . ①TU984-39

中国版本图书馆 CIP 数据核字(2018)第 199939 号

随着信息通信技术的迅猛发展,大数据已成为重要的科学研究方向,并在多个学科中发挥着积极作用。其中大数据带给城市规划和城市研究的影响尤为显著。笔者于2016年秋季学期在清华大学首次开设了“大数据与城市规划”研究生课程,通过理论与城市规划实践结合的方法进行授课。可作为高等学校城乡规划及相关专业教材,供本科生及研究生使用;也可为规划师、城市研究者等相关领域的技术及研究人员提供参考。欢迎各位读者随时将阅读或学习本教材时的反馈发送至 ylong@tsinghua.edu.cn。

责任编辑:杨虹 尤凯曦

责任校对:姜小莲

住房城乡建设部土建类学科专业“十三五”规划教材
高校城乡规划专业规划推荐教材

城市规划大数据理论与方法

龙瀛 毛其智 著

*

中国建筑工业出版社出版、发行(北京海淀三里河路9号)

各地新华书店、建筑书店经销

北京雅盈中佳图文设计公司制版

北京中科印刷有限公司印刷

*

开本:787×1092毫米 1/16 印张:22 字数:439千字

2019年1月第一版 2019年1月第一次印刷

定价:56.00元

ISBN 978-7-112-22592-7

(32698)

版权所有 翻印必究

如有印装质量问题,可寄本社退换

(邮政编码 100037)

前言

2015年11月10日，习近平总书记主持召开中央财经领导小组第十一次会议，强调要“做好城市工作，首先要认识、尊重、顺应城市发展规律，端正城市发展指导思想”。2017年12月9日，习近平总书记在中共中央政治局第二次集体学习时强调，“审时度势、精心谋划、超前布局、力争主动，实施国家大数据战略，加快建设数字中国”。总体上，城市大数据的出现和应用，为更客观地认识城市发展规律和更科学地支持城市规划的制定提供了极好的条件。

在大数据时代，进行城市研究需要处理海量、多源的时空数据，这对规划师提出了更高的技术要求。除了掌握城市规划的基础知识，规划师还需要具备包括计算机应用、数理统计等在内的一系列能力。英美部分知名高校（如麻省理工学院、伦敦大学学院和纽约大学等）已经开设了“城市模型”、“大数据与城市规划”以及“智慧城市”等相关课程。其所使用的教材多是授课教师的专著或最新研究论文的合集。考虑到大数据还是相对较新的概念，虽然其在我国城市规划中已经引起了较大的反响并积累了较多的应用经验，但国内尚未出版相关教材。随着我国城市化进程的转变，以及城市发展问题复杂性和综合性的增加，我国城市规划教育界已经意识到提高大数据专业教育水平的重要性。因而根据高等学校城乡规划学科专业指导委员会的要求，城乡规划专业的培养计划将增设较多定量城市研究内容，包括纳入数据统计分析、城市发展模型、地理信息系统、城市规划公众参与等诸多课程或知识点。

在这样的背景下，顺应我国城市规划编制和国内对城市规划教育变革的需求，笔者集成自身的国际化学术研究、本土化工程实践以及海内外学术交流经历，于

2016年秋季学期在清华大学首次开设了“大数据与城市规划”研究生课程，力争做到理论方法与城市规划实践结合，以期使之成为学生知识结构的重要内容，并提高学生在大数据分析与量化城市研究方面的能力。为了将该课程内容传播给更多城乡规划专业的学生及从业者，笔者将课程内容重新编撰集结为本教材。具体而言，本教材既可作为高等学校城乡规划及相关专业教材，供本科生及研究生使用；也可为规划师、城市研究者等相关领域的技术及研究人员提供参考。希望以此能够促进大数据在城市规划教学和实践中的应用不断深入。

大数据无论在国际还是在国内都是较为新鲜的事物，在城市规划领域更是仅仅只有几年的讨论与应用，因而对城市规划大数据的理论与方法框架进行组织的难度可想而知。因此本书难免有对已有讨论覆盖不全、技术方法不代表最新实践等诸多问题，这还有待于下一版本的修订。欢迎随时将阅读和使用反馈发送至 yulong@tsinghua.edu.cn。

此外，本教材使用者可以参加笔者在学堂在线（<http://www.xuetangx.com>）开设的MOOC大规模在线公开课《大数据与城市规划》进行进一步学习。详细课程网址为：http://www.xuetangx.com/courses/course-v1:TsinghuaX+70000662+2018_T2/about。

更多相关研究，请访问北京城市实验室网站（<http://www.beijingcitylab.com>）或个人网页（<http://www.beijingcitylab.com/longy>）。

龙瀛 毛其智

2018年5月于清华园

目录

- 001 第1章 概论
- 003 1.1 城市空间新数据环境
- 004 1.2 定量城市研究
- 005 1.3 新数据环境下定量城市研究的四个变革
- 008 1.4 大数据在城市研究与规划设计支持中的五个维度
- 009 1.5 代表性研究机构
- 010 1.6 本教材特点及结构
- 010 参考文献

- 011 第2章 变化中的中国城市与未来城市
- 013 2.1 过去与现在
- 015 2.2 当下城市发生的变化
- 023 2.3 未来城市的研究与规划设计
- 024 参考文献

- 025 第3章 城市新数据类型与典型数据介绍
- 027 3.1 传统数据
- 028 3.2 新数据环境：城市空间新数据
- 032 3.3 典型城市空间新数据介绍
- 044 3.4 目前新数据存在的问题及相关思考
- 046 参考文献

047	第4章 城市大数据的获取与清洗
049	4.1 空间大数据及获取
067	4.2 数据清洗
073	第5章 城市大数据统计与分析
075	5.1 大数据空间分析
087	5.2 大数据统计分析
104	5.3 其他分析与统计工具
109	第6章 城市大数据的可视化
112	6.1 基于 ArcGIS 的可视化
123	6.2 基于极海 (GeoHey) 的可视化
126	6.3 其他可视化工具
131	第7章 城市大数据挖掘：空间句法
133	7.1 基本介绍
134	7.2 空间句法与城市大数据
135	7.3 轴线分析与线段分析
140	7.4 空间句法软件 DepthMap
144	7.5 空间句法与大数据结合的应用： 北京街道活力——测度、影响因素与规划设计启示
150	参考文献
151	第8章 城市大数据挖掘：城市网络分析
153	8.1 城市网络分析概述
162	8.2 中国城市规划领域的知识网络
165	8.3 基于交通工具轨迹透视城市网络关系
170	8.4 城市社会空间网络联系
174	参考文献
177	第9章 大模型：跨越城市内与城市间尺度的大数据应用
179	9.1 大模型的提出背景
181	9.2 大模型：城市与区域研究的新范式
183	9.3 大模型为城市研究提供新的视角和思考
189	9.4 小结
190	参考文献

- 191 第 10 章 数据增强设计
- 193 10.1 在新数据环境下探索城市秩序的可持续内涵
- 195 10.2 数据增强设计 (DAD) 概念
- 195 10.3 DAD 体系结构
- 201 10.4 DAD 应用案例
- 208 10.5 国内外 DAD 相关研究机构、中心及团队
- 213 参考文献

- 215 第 11 章 基于公交卡大数据的城市空间研究
- 217 11.1 数据形式
- 221 11.2 数据处理
- 224 11.3 数据处理结果
- 230 11.4 数据的应用
- 234 参考文献

- 237 第 12 章 基于社交网络大数据的城市空间研究
- 239 12.1 微博数据
- 244 12.2 签到 (check-in) 数据
- 246 12.3 大众点评数据
- 252 参考文献

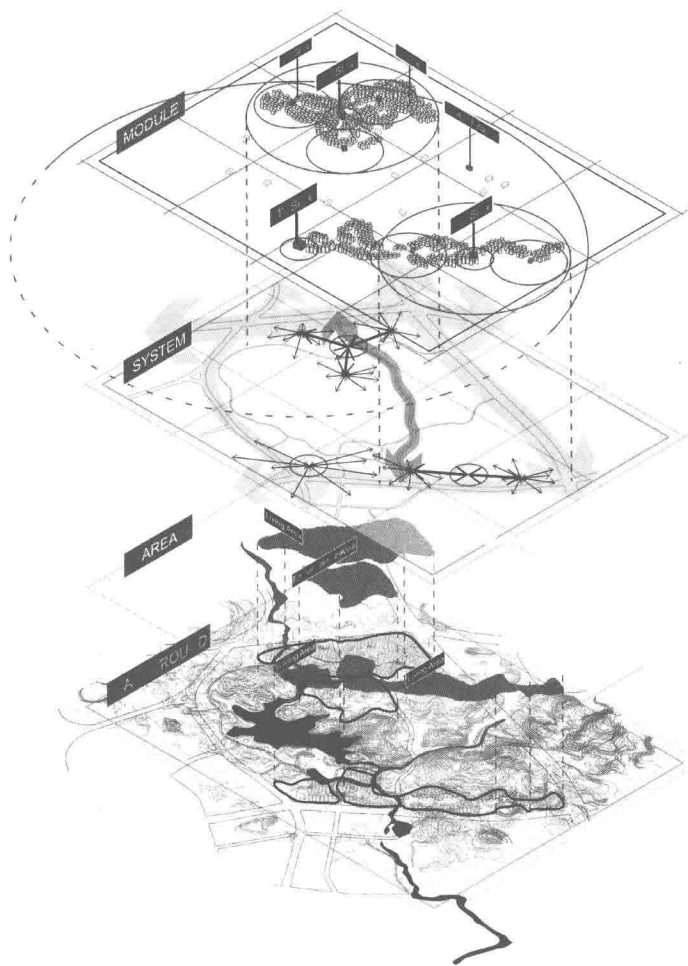
- 253 第 13 章 基于图片大数据的城市空间研究
- 255 13.1 图片大数据
- 256 13.2 图片来源
- 259 13.3 图片分析与可视化工具
- 261 13.4 相关研究
- 262 13.5 实践案例
- 271 参考文献

- 273 第 14 章 基于手机大数据的城市空间研究
- 275 14.1 数据介绍
- 276 14.2 数据获取及数据处理
- 279 14.3 基于手机大数据的城市空间研究
- 282 14.4 手机大数据应用案例
- 290 参考文献

- 291 第 15 章 总体规划中的大数据应用
- 293 15.1 城市总体规划的大数据应用框架
- 295 15.2 大数据在现状评估中的应用
- 301 15.3 大数据在预测未来中的应用
- 307 15.4 大数据在规划实施评估中的应用
- 313 参考文献
-
- 315 第 16 章 城市设计中的大数据应用
- 317 16.1 城市设计概述
- 319 16.2 基于大数据的城市设计方法
- 325 16.3 大数据在增量型城市设计中的应用
- 331 16.4 大数据在存量型城市设计中的应用
- 341 参考文献
-
- 342 后记

第1章

概论



近几年，随着信息通信技术的迅猛发展，大数据已成为重要的发展方向和研究领域，在多个学科都发挥着积极作用。相比其他传统行业，大数据带给城市规划和城市研究的影响更为显著，其不仅对城市规划编制、评价和管理的方式产生影响，也通过对人的活动、移动和交流方式的改变，改变了城市规划的对象——城市。大数据的应用与智慧城市理念促进了城市规划的科学化与城镇治理的高效化，使得各部门在数据及时获取与有效整合的基础上，能够及时发现问题，实时进行科学决策与响应；同时为公众参与提供了基础与平台，为以人为本、面向存量、自下而上的新型城市规划构建提供了基础。

本章介绍了当前时代背景下由大数据与开放数据共同构成的新数据环境，及其对定量城市研究带来的若干种变革，并列出了当前在大数据领域较有代表性的研究机构及团队。最后对本教材的特点及结构进行了介绍。

1.1 城市空间新数据环境

在过去，中国的城市规划和城市研究长期受制于数据的获取。这是由于规划工作的数据高度依赖官方的测绘数据、统计资料以及政府的行业主管部门的官方数据。而开放数据运动的出现正逐渐改变这个局面。开放数据（open data）指的是一种经过挑选与许可的数据，这些数据不受著作权、专利权以及其他管理机制所限制，可以开放给社会公众，任何人都可以自由出版使用。政务网站、商业网站（包括在线社交媒体）及科研网站构成了目前开放数据的三大重要来源。

而大数据起源于2000年以来信息通信技术（ICTs）的发展，尤其是社交媒体门户网站、智能手机应用的普及，以及云计算的诞生。实际上，大数据并不简单等同于数据量大的数据。目前较为广泛接受的大数据本质特征可概括为5个V——Volume（数据量）、Variety（多样性）、Velocity（速度）、Veracity（真实性）和 Value（价值）。大数据的种类多、价值高，但真正投入规划行业使用的却相对不丰富。目前规划业主要运用的大数据类型有公交卡数据、基于位置服务（Local-Based Service、LBS）的数据、浮动车数据及手机数据等。需要指出一个范畴上的误区：大数据的概念有广义与狭义之分。国内已有的多数研究，实际上多不属于狭义的“大数据”，而是属于“开放数据”（即广义的“大数据”），比如来自商业网站或政府网站的数据；而狭义的大数据如手机信令、公共交通刷卡记录和信用卡消费等记录大多不是开放数据，其获取难度大、成本高，因而目前存在着“大数据不开放，开放数据不大”的现象。这制约了规划实践及城市研究中对数据的获取与运用，与当前“开源与众包”的新时代精神理念有所背离。这是规划行业拥抱数据，谋求发展所面临的挑战之一。

狭义的大数据与开放数据共同形成了有别于传统调研和统计数据的新数据环境（new data environment），而这也是十年前并未广泛使用的数据。与传统数据相比，新数据环境主要呈现出精度高（以单个的人或设施为基本单元）、覆盖广（不受行政区域限制）、更新快（每月、每日，甚至每分钟更新）等特点。它不仅意味着更大的数据量，更反映了数据背后关于人群行为、移动、交流等维度的丰富信息。新数据环境日益成为国内城市规划学界、业界和决策界的共同关注热点，让学者、规划师和决策者观测到社会个体及详细空间单元上的丰富信息，为城市研究、规划设计、工程实践和商业咨询等带来了新的契机。

本教材中如无特别说明，均用“大数据”指代“新数据环境”。

1.2 定量城市研究

数据在城市规划与研究中的应用，多是在定量城市研究（quantitative urban studies）方向。它是指在一定理论基础之上，采用各种数据和技术方法，致力于探索城市发展的一般规律，并诊断城市问题、模拟城市运行、评估发展政策、寻求解决方案的科学研究方法，可应用于支持城市规划现状分析、方案编制与方案评估等各个阶段。当前的定量城市研究并非单纯的传统意义上的规划信息化或规划新技术的应用，而更注重对城市现象客观、直观、全面的分析，并通过多种媒介将分析结果传达至政府、规划师、专家学者和城市居民，从多方面提高城市规划与相关政策制定的科学性。

在国外相关研究领域，利用大数据开展城市空间与人群活动分析已成为当前学

术界的研究热点，所催生的大量研究可初步归纳为七种类型，包括社交网络数据的实时描绘（real time sensing）、多种交通网络数据分析（multiple networks）、城市新型数据系统构建（new urban data systems）、新型交通模型（new models of movement and location）、城市发展路径风险分析（risk analysis of development path）、新型人群移动分析系统（new models and systems for mobility behavior discovery），以及新型交通需求管理工具（new tools for governance of mobility demand）（Batty, et al. 2012）。相关研究呈现出三方面特点：首先是研究趋于片段化（fragmented），即对城市现象的某一个具体方面的局部分析多于对城市系统的综合分析，这与大数据数量大、精度高，但维度较少的特点有关；第二是分析算法趋于简单化，也就是通过简单的时间、空间和属性层面的统计分析，就可以得到有趣的分析结果；第三是更侧重对现状问题的识别和分析，而非对未来的预测或模拟。

目前，我国利用大数据开展的具有一定代表性的定量城市研究包括：利用公交刷卡记录研究通勤出行、城市贫困、过度通勤、公交通勤空间结构等问题；利用手机信令数据研究城市人口分布、空间结构、商圈影响力、居民出行距离等；利用出租车 GPS 数据预测拥堵地点；利用居民活动 GPS 数据分析城郊居民日常活动时空特征；利用社交网络位置数据和签到信息研究城市用地功能与混合度、城市发展边界、城市活动区域划分、城市网络信息空间结构；利用百度指数研究区域城市网络特征；利用百度、高德迁徙数据研究城镇体系、居民黄金周旅游行为等；利用大众点评数据研究餐饮业格局、餐厅选址等；利用全国 PM2.5 监测的在线数据研究 PM2.5 污染分布；利用微观尺度的人口统计数据分析中国城镇格局等。城市大数据的出现与成熟带来了城市研究领域关注内容和研究方法的较大变化，促使城市规划与其他相关学科的进一步融合，以及在研究范式、研究方法与内容上的革新。

1.3 新数据环境下定量城市研究的四个变革

在我国城市规划逐渐由过去二三十年的“大拆大建”向精细化规划编制与管理转型的背景下，定量城市研究得到了越来越多的关注。同时，近年来大数据的涌现和一定程度上的普及，为定量城市研究提供了大量新的数据来源。其中既包括严格意义上的大数据，也包括多源的开放数据。两者构成了了解城市系统运行规律的重要基础，并推动了定量城市研究领域在多个方面的巨大变化。

1.3.1 空间尺度的变革：从小范围高精度、大范围低精度到大范围高精度

在传统数据环境下，受数据收集方法的限制，城市和区域研究在研究覆盖范围和精细度上往往很难做到两者兼顾——大范围的研究通常以牺牲精细度为代价，而

精细度高的研究往往覆盖范围较小。因此，传统数据环境下的城市和区域定量分析可主要分为两种：一是针对单一城市做较为深入的研究，如研究城市贫困问题、公共服务设施的配套水平等；二是覆盖全国或多个省市地区的区域分析，如一些宏观经济研究，多以县、市或省为单元，研究单元较大，难以反映小尺度信息。

而新数据环境为在较大空间范围内收集高精度数据提供了可能，如社交网络和各类商业网站数据往往覆盖全国且以人、车、商户等个体为基本单位，可充分满足精细化的分析需求。而对某些传统数据的有效整合也有利于拓展数据的广度与精度。如以往人口密度研究主要在区县尺度，属于宏观分析的范畴，在新数据环境下则可将研究尺度缩小至乡镇街道级别。这不仅促进了研究范围的扩大和精度的提高，且有助于呈现以往难以发现的新的问题。国外相关研究包括：通过利用通话记录重新划定城市和区域范围；利用带地理标签的推特（Twitter）数据进行人群活动类型识别，并结合多主体建模方法构建城市动态模型；利用高精度的全国人口调查数据研究人口聚集规模等。

针对这一数据特点，笔者提出了“大模型”研究范式，并试图通过这一范式在城市模型研究中兼顾覆盖区域乃至全国的研究范围与精细化的城市模拟单元（详见第9章）。随着新数据环境的不断成熟，传统城市研究的“研究地盘”的概念也将会逐渐弱化，学者有望对千里之外的城市进行深入的城市研究。

在传统数据环境下，中小城市和县、镇与大城市之间在数据基础和研究水平方面存在明显差距——城市研究多关注具有代表性的大城市而忽略了二、三线城市和等级更低的城市。“大模型”通过关注全国绝大多数城市，有望在一定程度上对中小城市发展给予更多研究关注、削弱技术差异，并系统探讨国家和区域城市化进程中各类城市的互动关系。此外，大范围、高精度的研究趋势有利于探索我国城市发展的一般规律。目前我国共有六百多座城市，不同学者针对不同城市采用不同方法开展的研究往往难以进行横向比较，而以大模型的为代表的新的研究范式可为一致性研究提供可能，从而在大样本城市研究的基础上探索我国城市发展的一般性和特殊性规律。

1.3.2 时间尺度的变革：从静态截面到动态连续

新数据环境所提供的另一重要突破是体现了不同时间尺度上的城市动态。传统城市研究的数据多来源于政府部门统计年鉴或抽样调查，数据以静态数据为主，只能反映某一时刻或一段时间内城市所处的状态（如年鉴对应一年、出行调查多对应一日），且由于数据取样的局限性，只能覆盖有限的空间范围。而包括公交刷卡、出租车轨迹、信用卡交易记录、在线点评以及位置微博和照片等在内的新数据环境，则可以反映个人乃至整个城市短至每秒、长至多年的动态变化，且具有连续性高、覆盖面广、信息全面等优势。例如利用精确到秒的信用卡交易记录，可以对城市每

小时的销售情况进行可视化,进而识别商圈;积累多年的信用卡交易记录,则可以体现出人们生活与消费方式的变化——如传统书店的萎缩和在线购物的繁荣。假设一千年之后的考古研究,能够发现人类的电子记录和足迹,则有望超越目前的考古发现,对此时的人类社会进行更为全面的重现。国外对于动态时间的相关研究已大量展开,如通过手机信号位置轨迹发现人们出行行为的规律性;通过手机信号记录研究城市的空间结构和多中心性;麻省理工学院感知城市实验室(MIT Senseable City Lab)与哥本哈根市合作,为五千块垃圾贴附地理标签,并在三个月内追踪垃圾流向,分析垃圾回收效率等。

1.3.3 研究粒度的变革:从以地为本到以人为本

新数据环境所提供的并非仅包括扩大的数据量,还包括数据所反映的城市居民的行为特征与规律,以及人对建成环境的感觉、情感、经验、体验、信仰、价值判断等。这些以前难以量化的因素在新数据环境中都可以得到有效的表达与数理分析。国外相关研究包括:利用多日手机数据挖掘城市人群活动不同的基本模式;利用 Flickr 图片数据分析人对城市空间的认知图像,从而重新阐释凯文·林奇的城市意象理论等。在我国新型城镇化的背景下,具有高粒度特性的新数据环境也为以人为本的城市研究提供了极佳的素材。在此数据平台基础上,居民行为、活动及其影响下的城市空间组织和结构的变化,社会群体特征、网络、活动等及其影响下的社会空间分异或融合等课题都可得到深入分析。

1.3.4 研究方法的变革:从单一团队到开源众包

众包(crowdsourcing)是互联网带来的新的生产组织形式,即利用互联网将原先单一机构内的工作任务以自由自愿的形式分配给机构外的志愿人员(通常为个人)完成,这一组织方式可以充分利用志愿者的创意和技能,以更低的成本、更高的效率完成任务。虽然开源、众包等概念听来与城市研究和城市规划领域相距甚远,但近两年来随着数据的开放和北京城市实验室^①(Beijing City Lab、BCL)等开放研究平台的成熟,众包模式也在逐渐融入定量城市研究和相关数据平台构建中(如搜集数据、学术合作、验证研究成果),并体现出优势。这种众包的城市研究方式有望突破传统的单一团队开展研究工作的模式,例如针对中国大量存在的收缩城市现象,探讨其背后的深层原因则需要大量实地调查,相关工作量远远超出了课题组或单一机构的承担能力,为此 BCL 于 2014 年 11 月发起了“中国收缩城市研究网络^②”,持续

① 北京城市实验室 <http://www.beijingcitylab.com>

② 中国收缩城市研究网络 <http://www.beijingcitylab.com/projects-1/15-shrinking-cities/>

跟踪收缩城市方面的国际研究并开展国内的理论和实证研究工作。总体上,在新数据环境下,城市研究的工作方法正在由单一团队向开源众包模式转变。

1.4 大数据在城市研究与规划设计支持中的五个维度

在城市研究与规划设计的过程中,大数据的分析与应用可以在不同阶段和不同层次进行辅助。基于城市开发周期理论,并结合中国城市的现实特点,可以从城市开发、城市形态、城市功能、城市活动、城市活力/品质/文化/风貌/特色等五个阶段/维度对城市系统进行监测、评价、情景分析乃至预测。

“城市开发 (development)”是指城市用地由其他用地(如农田、农村建设用地)转变为城镇建设用地。应用于测度城市开发维度的数据包括遥感解译的土地利用情况、用地现状图、房地产数据等。

“城市形态 (morphology)”是指城市用地布局、路网分布等。应用于测度城市形态维度的数据包括分等级路网、道路交叉口、建筑物、土地出让/规划许可、街景图片等。

“城市功能 (function)”是指居住、就业、交通、游憩等城市用途在空间上的组织(包括密度)。应用于测度城市功能维度的数据包括地块尺度的用地性质、兴趣点^①(POI)等。

“城市活动 (activity)”是指居民在城市内部从事的各种类型的活动(如工作、出行、休闲、社交等)。应用于测度城市活动维度的数据包括企业数据、手机数据、微博和大众点评等社交网络数据、公交卡数据、百度热力图等。

“城市活力 (vitality)”是指城市活动所具备的内在特征(如品质、文化等)。应用于测度城市活力维度的数据包括点评类数据、手机数据、位置照片、房价信息等。

这五个维度是空间递进关系,只有农村用地转为城镇用地,才能够通过规划建设实现相应的城市形态;有了纯粹物质空间的城市形态,才能够造就相应的城市功能;有了城市功能,才能够在社会空间促进相应城市活动的产生;而有了城市活动,才可能体现出良好的城市活力、品质、文化等。这几个维度依次递进,有开发没有形态,有形态没有功能,有功能没有活动,有活动没有活力,都属于目前中国大量城市的突出问题。大量中国城市在空间维度存在的问题,可以用相互链接的两个维度进行刻画和解释。

新数据环境为从精细化尺度评价城市开发、城市形态、城市功能、城市活动和城市活力提供了前所未有的机遇。这五个维度构成了城市大数据支持城市规划编制

① 兴趣点 Points of Interest