

新数据环境下的城市研究、规划与设计

龙瀛，博士，高级工程师

北京市城市规划设计研究院，北京市南礼士路 60 号，100045，北京

1 误区与契机

信息通讯技术的发展以及政务公开的推进使大量数据如雨后春笋般涌现，人们称之为大数据。需要指出一个范畴上的误区是，大数据的概念有广义与狭义之分，国内已有的多数研究（包括我本人的研究），实际上都不属于大数据，而是属于开放数据，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，获取难度较大。为此，本讨论用“新数据环境”作为目前大数据、开放数据等研究工作的统称。

新数据环境日益成为国内规划学界和工程学界的关注热点，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活动。新数据环境产生的微观个体（社会层面和物理空间层面）数据，与“以人为本”的新型城镇化不谋而合，对存量规划、收缩城市、公共参与等热点问题具有支撑作用，为城市研究、规划设计、工程实践和商业咨询等带来了新的发展契机。

新数据环境下，国内多个城市研究和规划机构开展了定量城市研究工作，如北京城市实验室的多数研究工作都是基于新数据开展的。新数据多以精细化尺度覆盖大量城市，为此我与合作者提出了“大模型”这种区域和城市研究新范式，并认为当前定量城市研究正经历四个方面的变革：空间尺度、时间维度、研究粒度和研究方法——这些变革都以提高居民生活质量为核心目的。

以下是我对新数据环境下研究工作的几个看法。

2 数据有偏性的应对

新数据的有偏性广为学界讨论。例如利用位置微博数据进行的城市幸福感研究，面临着多个层面的偏差，包括微博用户的口是心非、自然语言处理的能力约束、发微博人群的代表性、新浪提供微博 API 的黑箱。这些偏差使得研究结果的可信性受到质疑。

针对这个问题，不同的研究采取了不同的应对方法，主要包括：

(1) 充分利用数据的有偏性。例如在中国城市中频繁乘坐公交的出行者中，低收入人群有更大的出现概率，我与合作者充分利用公共交通刷卡数据（smart card data/SCD）的有偏性，研究城市贫困对象在 2008-2010 年间的出行行为、居住地和就业地的变化情况；

(2) 关注局部群体的行为特征。我与合作者利用学生 SCD 分析北京 985 高校学生的出行行为，也分析了四类极端出行人群的出行行为及其社会经济特征；

(3) 结合其他数据研究同一问题，以保证研究结果的稳定性。如我与合作者利用 SCD、传统居民出行调查数据、社交网站签到数据以及出租车轨迹数据共同评价北京城市规划的实施效果，发现各类数据的评价结果都表明超过 95% 的人群活动与规划边界一致；

(4) 在承认新数据有偏性的前提下研究城市系统，并期待与其他有偏数据研究互补，一起勾勒出完整的城市系统。

3 多瞬时的可视化，少长周期探索

已有的新数据研究多利用少于一周时长的记录，部分研究仅处于可视化层次，而少有利用覆盖多年的数据进行的研究。与新数据相比，传统数据如城市年鉴能反映出城市一年的变化和年末状态，以及城市的多年累积变化。随着新数据的不断积累，多年连续的新数据研究成果和规划应用，将与目前已发表的研究结果有很大差异。例如，一日信用卡刷卡数据，能够识别城市的消费规律，一个月的刷卡数据，可以识别节日的影响，多年的数据则能够看到技术进步对消费行为的改变。这样的研究，也将突破现有研究被质疑多可视化、多现状、少理论的瓶颈。我和合作者利用 2008-2014 的 SCD 开展了若干探索性研究。

4 多现状分析，少面向未来的规划设计支持

已有的研究多针对城市系统的现状评价和问题识别，而少有面向未来的研究。新数据环境下开展的诸多研究的成果，需要适时反哺面向未来的规划设计。以往的规划支持系统并没有有效地支持规划设计，为此我与合作者提出了数据增强设计(Data Augmented Design/DAD)的规划设计新方法论。

DAD 是以定量城市分析为驱动的规划设计方法，通过数据分析、建模、预测等手段，为规划设计的全过程提供调研、分析、方案设计、评价、追踪等支持工具，以数据实证提高设计的科学性，并激发规划设计人员的创造力。DAD 的定位是现有规划设计体系下的一种新的规划设计方法论：它不是艺术设计的背叛者，而是强调定量分析的启发式作用的一种设计方法，致力于减轻设计师的负担而使其专注于创造本身，同时增强结果的可预测性和可评估性。DAD 利用简单直接的方法，充分利用传统数据和新数据，强化规划设计中方案生成或评估的某个环节，易于推广到大量场地，同时兼顾场地的独特性。近期我和合作者将发表 DAD 的相关文章。