

主旨
聚焦

风起云涌的“大数据”热潮无可避免地波及到地理学领域,地理学碰上“大数据”是一次新的发展机遇与挑战,还是一阵风似的热潮与炒作?本次笔谈将聚焦这一论题展开不同观点的碰撞。

地理学碰上“大数据”:热反应与冷思考

吴志峰¹,柴彦威²,党安荣³,龚建华⁴,高松⁵,乐阳⁶,李栋⁷,
柳林⁸,刘行健⁹,刘瑜¹⁰,龙瀛¹¹,陆锋¹²,秦承志¹²,
王慧¹³,王鹏¹⁴,王伟¹⁵,甄峰¹⁶

(1. 广州大学地理科学学院,广州 510006; 2. 北京大学城市与环境学院,北京 100871;
3. 清华大学建筑学院,北京 100084; 4. 中国科学院遥感与数字地球研究所,北京 100094;
5. 美国加州大学圣塔芭芭拉分校,美国加州圣塔芭芭拉 93106; 6. 深圳大学土木工程学院,深圳 518060;
7. 中国城市规划设计研究院,北京 100044; 8. 中山大学地理科学与规划学院,广州 510275;
9. 香港城市规划与设计系,香港; 10. 北京大学地球与空间科学学院,北京 100871; 11. 北京市城市规划设计研究院,北京 100045; 12. 中国科学院地理科学与资源研究所,北京 100101;
13. 厦门大学建筑与土木工程学院,厦门 361005; 14. 北京清华同衡规划设计研究院有限公司,北京 100085; 15. 中央财经大学政府管理学院,北京 100081; 16. 南京大学建筑与城市规划学院,南京 210093)

摘要: 互联网时代的“大数据”热潮迅猛波及到经济社会的各个领域,地理学是大数据研究与应用的天然试验场。聚焦地理学与大数据的碰撞,回顾大数据在地理学研究中的应用探索,重点讨论大数据给地理学研究与发展带来的机遇与挑战。讨论认为:大数据已经对地理学研究产生了一定的影响。其中,人文地理学领域的反应最为热烈,基于大数据的研究案例纷纷呈现;地理信息科学在互联网大数据时代将会更加迅猛发展,自然地理学领域正在寻找和等待爆点。目前,大数据还不能改变地理学的核心命题与基本范式,在坚持地理学核心思想的同时,应该对地理学领域有关大数据理论的探讨与应用尝试持有一种开放包容的态度。

关键词: 地理学;大数据;机遇与挑战;笔谈;反思

DOI: 10.11821/dlj201512001

主题解析

地理学作为一门古老的学科一直顽强地延续与发展着,在经历各种危机后始终保持着旺盛的生命力。但是也不可否认,当今地理学学科内部呈现一种明显的破碎化、离心化趋势。有学者批评当今地理学:自然地理破碎化、人文地理经济化与社会化、地理信息科学信息技术(information technology, IT)化,这个说法虽然有失偏颇,但也不无道理。地理学“区域性”与“综合性”的传统学科特色与优势遭到忽视与弱化,然而当今地理学面对的科学与社会问题越来越复杂,地球表层系统作为地理学研究的核心对象,受人类活动干扰与改造的影响越来越显著,因此,跨学科的综合与协同研究成为地理学

收稿日期: 2015-06-12; 修订日期: 2015-11-03

基金项目: 国家自然科学基金项目(41171446); 广州市市属高校“羊城学者”科研项目(12A002G)

作者简介: 吴志峰(1969-),男,湖南湘潭人,博士,研究员,博士生导师,主要从事城市遥感、景观生态、城市化复杂地表过程方面的研究。Email: gzuwzf@163.com

学科发展的必然趋势。

与其他自然或社会科学一样,地理学的发展必须解决两个问题:一是独特与关键科学问题的引领。美国国家科学委员会在2010年抛出了未来10年美国地理学的11个前沿战略方向^[1,2];傅伯杰院士在中国地理学会第十一次全国会员代表大会报告集思广益,也提出了中国地理学从地理学走向地理科学所面临的重大命题,当今地理学或者地理科学面临的科学问题是明确且具有重大意义的,这决定了地理学的不可替代性^[3]。二是解决科学问题的技术手段与方法途径。由于地理学研究对象的复杂性、跨自然与人文的综合性,在研究方法、技术手段与研究范式等方面,地理学相对落后,这也成为制约地理学快速发展的一个不可忽视的因素。

“大数据”是继“云计算”之后互联网时代掀起的又一个热潮。与“云计算”相比,“大数据”具有更强的张力。当云计算还掌握在少数机构与人群手中,高挂在云里雾里的时候,大数据已从IT领域迅速向外扩散,在商界、政务管理、各学科领域乃至社会民生中引起了高度的关注与热烈的反响。在大数据理论体系尚未成型之时,大数据技术、方法与应用实践已走在前面。在互联网时代,“大数据”已经成为人类社会的一种重要资源,而且这种资源是呈几何级数般的增长,并且迅猛地影响着当今人类社会的发展与转型。2015年10月4日,国务院正式印发了《促进大数据发展行动纲要》。这一行动纲要的出台意味着大数据发展正式成为国家战略。

有学者认为“大数据”的产生是科学研究“第四范式”的出现,现在下这样的结论还为时过早,但已经不可避免地对各学科产生冲击。当地理学碰上“大数据”后,同样产生一股热潮,并且学科内部反应的热度不一:人文地理学反应最为热烈,城市地理、社会地理、行为地理、商业地理、旅游地理、交通地理等基于大数据的研究案例纷纷呈现,令人耳目一新;地理信息科学(遥感、综合对地观测、测绘等)则在采集、管理与应用“地理大数据”方面走在前面,海量的遥感数据、对地观测数据、测绘地理信息数据本身就是大数据的重要组成;自然地理学领域则显得冷相对静,也许对大数据的渴望没有人文地理研究那么强烈,也许大数据与自然地理研究的结合点与爆发点还在酝酿中。

当前大数据热潮的冲击是否会对地理学的发展带来挑战与机遇?地理学各分支学科和学科方向对大数据热有何不同的反应?其原因是什么?地理学计量革命与地理信息革命之后,大数据是否对地理学理论发展与研究范式产生深远影响?抑或大数据本身就是一阵概念的炒作,过后就会销声匿迹?地理大数据的有何独特之处?如何顺应“互联网+”与大数据的时代潮流,推动学科的发展,将地理学研究与社会经济发展紧密结合起来?本次笔谈邀请到10多位不同年龄层次、不同学科背景的学者展开思想与观点的碰撞,提出的见解仅代表笔谈作者本人,并不代表本刊的立场或观点。

主题主持人及其概况:



吴志峰(1969-),男,湖南湘潭人,博士,研究员,博士生导师。1992年毕业于湖南师范大学地理系,1995年于华南师范大学获自然地理学硕士学位,2002年于中科院地理资源所获地图学与GIS博士学位。现就职于广州大学地理科学学院,担任中国地理学会理事、国际景观生态学会中国分会副理事长、广东省地理学会副理事长、广东省遥感与GIS学会副理事长等职务。主持国家自然科学基金项目2项,其他省部级以上课题10余项。近年来主要从事城市遥感、景观生态、城市化复杂地表过程方面的研究。Email: gzuwzf@163.com

主题主持人观点

• “大数据”必然会对地理学的发展产生影响，但是把它看作是一次革命性变化尚为时过早。我们抱着一种开放与包容的态度去接纳它，开展更多的实践应用与理论探讨，既不能拒之千里，也不应该将其看作包治百病的灵丹妙药，“大数据分析”替代不了地理分析与地理研究。

• 大数据不单纯是“数据”，更是一种能力和资源，大数据将改变很多传统的生活、工作、管理和思维模式。大数据之“大”，不仅在于其规模容量之大，更在于人类可以处理、分析并使用的数据在大量增加，通过这些数据的处理、整合和分析，可以发现新规律，获取新知识，创造新价值。

• 大数据与“小数据”是否存在明显的界限？具有时空特征的“大数据”（地理大数据）是否是“大数据”的主流？可以肯定的是，失去地理位置（或者空间信息）的“大数据”以及没有时间维度的“大数据”将成为死数据。

• 大数据的“富矿”在哪里？“城市”是一个“人”高度聚集的空间区域，通过多源“城市感知”获取的“城市大数据”是宝贵资源，如交通大数据、市政大数据等，由于人文地理与城市规划对城市与“人”的行为空间特征更为关注，因而对大数据应用反响更热烈。

• 数据只有得到应用才能增值，才具有意义。阻碍当前大数据应用与作用发挥的主要障碍有哪些？城市管理中如何更智慧？城市与政务大数据的潜在价值？自然地理与生态环境研究的“大数据”何在？遥感与综合对地观测获取的大数据如何更好地应用于复杂地表过程与人居环境研究？这些都值得思考。

各位嘉宾观点观点（按姓名拼音字母排序）



柴彦威

（北京大学城市与环境学院，博士，教授；中国地理学会城市地理专业委员会主任委员）

观点：“大数据”的兴起与迅猛发展有利于创新地理学的研究方法与研究议题；“大数据”不是万能的，各类数据扬长避短、各得其所，隐藏在数据背后的事实将被更多地揭示出来；地理学界对“大数据”应予以更多关注。

人文地理研究不仅需要利用新的数据来源对已有理论和模型加以再验证，而且更需要思考大数据带来的新机遇、新认识、新方向。

首先，大数据的应用符合当前学科“人本化”与“信息化”的学科发展趋势。大数据提供了大量关于“人”的数据，使人地关系研究更能够关注“人”，更好地体现以人为本。其次，大数据能够在一定程度上克服传统数据的不足。传统数据如问卷调查、访谈等能够获取的往往是总体样本的一个子集，存在样本代表性和典型性等问题，对于地理空间格局与过程的理解也因数据限制而存在偏差，大数据的相关关系分析与传统数据的因果关系研究可以相得益彰。再次，大数据能够为传统研究议题提供新的佐证。例如在通勤研究中，利用不同的调查数据，针对同一城市的研究结果可能存在较大差异；相应

地,对于机制的解释也会有所不同。大数据作为一种“完全数据”,为整合和评判不同研究结论及其社会意义提供了一种可能性。最后,大数据也为创新地理学研究的议题提供了机遇。例如在微观尺度上对于个体和群体移动模式的预测,在宏观尺度上应用多维度行为指标可提升都市区研究中对于人与空间互动关系的理解。

“大数据”不是万能的,但各类数据都有其价值,只有扬长避短,各得其所,隐藏在数据背后的事实将被更多地揭示出来。呼唤地理学界对“大数据时代”给予更多的关注。



党安荣

(清华大学建筑学院,博士,教授;清华大学人居环境信息实验室主任;中国城市规划学会新技术应用学术委员会副主任委员;中国地理信息产业协会教育与科普工作委员会副主任委员)

观点: 大数据可动态揭示人的空间活动与空间行为模式,为城市地理学的研究提供动态的空间组织依据,促进城市地理学研究由静态空间转向动态空间、由土地空间转向行为空间。

城市地理学是研究城市(镇)的形成、发展、空间结构和分布规律的学科。其核心研究内容是从区域的空间组织和城市(镇)内部的空间组织两种地域系统考察城市(镇)的空间组织,而这种空间组织的根本依据就是人的空间活动与空间行为模式。

大数据多样性、真实性、代表性、动态性等特点,可以分别从宏观、中观、微观等层次,动态揭示人的空间活动与空间行为模式的规律,为城市地理学研究提供动态的空间组织依据,促进城市地理学研究由静态空间转向动态空间、由土地空间转向行为空间。

具体而言,在宏观层面(全球、区域):社交网络大数据、网络搜索大数据、移动通信大数据、航天遥感大数据等的分析与应用,可以动态揭示区域城镇化乃至全球城镇化的动态发展,为城镇体系及城镇群的规划与发展研究提供科学依据。在中观层面(城市、城镇):网络签到大数据、定位导航大数据、交通智能卡大数据等的分析与应用,可以揭示城市(镇)居民的时空分布与时空行为模式,反映城市用地及其功能布局、城市交通及其道路设施的合理性,为城市(镇)总体规划与发展研究提供科学依据。在微观层面(社区、建筑):居住环境感知大数据、建筑与生活能耗大数据、室内定位及居民行为大数据、精细化网格管理大数据等的分析与应用,可以揭示居民在社区乃至居室的行为模式,为城市(镇)详细规划、社区管理服务及建筑设计提供决策支持。



龚建华

(中国科学院遥感与数字地球研究所,博士,研究员;中国计算机学会虚拟现实与可视化技术委员会委员)

观点: 大数据对于地理信息科学的发展是一个巨大的新机遇,是否为地理系统的复杂性科学研究提供了新的思路与方法?

2004年,与林琚教授从虚拟地理环境及SARS等疾病时空传播的研究角度,提出面向“人”的地理信息系统(GIS)概念,认为传统的GIS侧重于“地”的表达与分析,是面向“地”的GIS,对区域人地系统中的“人”,即个体、群体、组织等研究很少,重视

不够。随着大数据时代的到来，尤其是关于“人”社会感知（social sensing）数据技术的发展，手机定位、车载GPS、在线地图导航与位置搜索、社交网络、城市感知高清视频等实实在在的新时空大数据涌现，极大地推进了面向“人”的地理信息系统（科学）的发展。但是，上述主要是从“人”这个具体研究对象以及新数据获取的角度，讨论关于地理时空大数据的影响。从更高的层面看，大数据是否属于科学研究的“第四范式”？大数据是否是重视“相关”关系而弱化“因果”关系研究？大数据是否为突破还原论、发展整体论与系统论提供了可能？是否为地理系统的复杂性科学研究提供了新的思路与方法？关于这些问题的提出、思考、激烈争论与跨学科讨论，呈现出的多种新观点、新思想以及相互激励碰撞形成的百家争鸣新气象，说明了大数据对学科发展的巨大影响力。

地理信息科学概念自从1992年提出以后，其基本理论与方法发展缓慢、创新活跃度不大，而大数据对于地理信息科学的发展则是一个巨大的新机遇，尤其在地理信息科学基本理论与方法论上将会不断呈现新思想、新概念与新方法。例如，已经出现了的开放GIS、后GIS世界、社会感知、面向“人”GIS、大数据GIS、广义GIS、城市计算等新观点；同时，结合大数据时代的虚拟地理环境、地理信息科学与地理科学发展，也提出了一些理论新思考，即地理大数据主要侧重于地理属性以及地理虚体（多种地理属性之间的确定性或必然性联系）的感知与表达，以及采用虚实二象思维研究现实地理环境、（实时）虚拟地理环境以及地理大数据世界的相互关系与影响等。



高松

（美国加州大学圣塔芭芭拉分校，地理信息学科博士研究生）

观点：空间大数据能够支持“人地交互”的传统地理学研究，并赋予了新的视角；大数据支撑下的地理学相关研究需要考虑结果的可验证性、尺度依赖性、不确定性、因果关系而不只是相关关系。

“空间”（space）与“场所”（place）是地理学的两大传统研究对象。在传统地理学的二级学科中，自然地理、地图学与地理信息系统更重视关于空间的研究，而人文地理更偏重于对场所的研究。大数据的出现，尤其是基于移动定位感知设备（GPS、手机、传感芯片等）收集的海量关于人的时空行为数据及社交媒体（博客、微博、互联网地图等）的用户行为数据，使得研究人员可以打破传统空间与场所的界限，通过研究地理空间中的大量人的移动行为轨迹和周边场所特征来预测人的活动规律、区域单元之间的交互特征，获取人们对场所的情感评价因素、评估动态变化的城市土地利用价值等。因此，空间大数据能够支持“人地交互”的传统地理学研究，并赋予了新的视角。

大数据驱动下的另外一个重要影响就是地理学研究范式的变化趋势。传统的地理学研究主要依赖于“自上向下”的理论建模驱动调查实验的设计，如今逐渐涌现“自下向上”的大数据驱动下的观察实证的研究。两种研究各有其优势和不足。调查设计类对样本的采样质量控制更好，人文社会经济属性变量丰富齐全，但是样本量相对较少、成本大、周期长；大数据驱动下的观察实证和知识发现，样本量大，成本低、周期短，但数据样本质量参差不齐，需要大量的数据预处理、机器学习和数据挖掘技术作支撑。空间大数据支撑下的地理学相关研究需要考虑结果可验证性、尺度变化性、不确定性、因果关系而不只是相关关系等重要因素。



乐阳

(深圳大学, 博士, 副教授; 中国GIS协会理论与方法专业委员会委员)

观点: 与经过严谨科学实验得到的数据相比, 绝大多数“大数据”是有偏数据; 解读大数据需要更严谨精密和鲁棒的算法; 大数据为研究个体、个体与个体, 以及个体与世界之间大规模的互动提供了新的方式, 地理学在这里大有可为。

从2008年开始接触出租车GPS数据, 后来逐渐接触环形线圈、手机, 以及公交卡等数据。在“大数据”这个名词出现前, GIS领域对数据的修饰定语一直用的是“海量、多源、异构”, 所以数据量大不是“大数据”最主要的特征。且数据量和计算能力的矛盾从计算机出现时就已经存在, 这两者的矛盾也一直是计算机硬件和软件发展的驱动力。

大数据之所以引发热议和思考, 更多是因为基于大数据产生的多角度、多层次、多粒度的分析或挖掘结果。几个大数据案例忽如一夜春风唤醒了大家的意识, 许多之前从来没有想到过的数据被发现、被产生, 并且被以各个方式应用到了从空气、水, 到城市、交通, 以及个体行为等的研究中, 改变了传统定量研究中缓慢、小范围的研究节奏。所以大数据带来了研究范式的变革^[4], 引发了震荡。

有一点必须认识到, 虽然现在大数据如火如荼, 不少政府相关部门也在积极推进大数据的运用, 但多数成果仍然是探索性的。例如谷歌流感趋势预测失效的案例^[5], 让大家开始反思大数据的使用。与经过严谨科学试验得到的数据相比, 绝大多数“大数据”是有偏数据(如手机数据倾向于电话使用频繁的人群, 微博数据倾向于年轻的群体); 易受多种因素影响和干扰。所以, 解读大数据需要更严谨精密和鲁棒的算法。在当前数据分析和挖掘算法水平下, 将大数据运用到专业领域, 仍然有一段比较长的路要走。传统定量研究方法需要, 也仍然会继续存在, 而大数据更适合作为传统方法的补充。

总之, 大数据为研究个体、个体与个体, 以及个体与世界之间大规模的互动提供了新的方式, 地理学在这里大有可为。同时, 数据管理和处理的手段继续以摩尔定律发展, 已经使得我们可以利用多样化工具不间断采集各类数据、建立系统化工具和设施。

用或不用, 大数据就在那里, 只增不减。



李栋

(中国城市规划设计研究院, 博士, 高级工程师)

观点: “大数据”中蕴含丰富的时空信息, 是地理学科最好的实验素材; 当前绝大部分的大数据生产者或拥有者与地理学的交集都很小; “大数据”在地理研究处于试验阶段, 不切实际的热捧或贬低都是有害的。

通过近几年的尝试与实践, 面对“大数据”热潮, 地理学领域所呈现的状态, 用四个“一”来加以概括。

“一点尴尬”: 数据可获得性并不确定, 直接限制了相关研究的开展。由于极少能直接产生大数据, 地理学者必须尝试通过各种途径去获得此类数据, 包括但不限于合作、共享、交换、购买、抓取等, 在分析开展之前已付出了大量的时间和精力成本。

“一丝窘迫”：地理学者对收集来的大数据的质量往往无法进行系统性评估，其中包括原始样本的偏差、预处理方法是否恰当，等等。大数据作为新生事物缺乏大量交叉校核的研究案例可供借鉴，导致分析和结论在可信及置信程度上都大打折扣。例如最近的工作发现，应用2015年的春运迁徙数据进行空间分析，与腾讯的结果相比，广为人知的百度迁徙数据其实存在较大的问题，但现实中大量的研究和规划应用已经在引用百度迁徙作为流动人口分析的依据和佐证。

“一些想法”：面对困境，地理学者能够做的依然是扎根现实、回归研究本质，冷静客观地评估各种“大数据”的适用条件，并依托学科领域及行业知识开展实证，尤其是基于多源数据相互检验等质量评估工作，一步步夯实大数据在地理中的应用基础。

“一份信心”：地理学科及城市规划等行业应用中，新数据的涌现、技术工具的革新已经成为确认的趋势和主要驱动力，基于复杂、动态、全面和细节导向的新研究方式正在并将替代单一、静态、断面和聚合导向的传统研究方式。数据正是这一切的催化剂。

社会感知理论是“大数据”在地理研究中的重要探索，但若类比已发展了半个世纪的遥感，这一方向依然处于襁褓期，任何不切实际的热捧或贬低都会带来不必要的伤害，只有耐心、严谨并负有责任去使用和推动才是唯一正确的方式。



柳林

(中山大学地理科学与规划学院，美国辛辛那提大学地理系，博士，教授)

观点：大数据作为一种新兴的科学范式以及一种新兴的数据源，对地理学发展将起到积极作用；过度热捧不利于大数据的持续发展和有效应用；目前国内谈大数据概念的多，用大数据解决实际问题的少。

目前国内谈大数据概念的多，期待看到更多的地理大数据应用实例。以犯罪地理研究为例，大数据时代之前很难获取受害人、犯罪分子以及公安人员的轨迹数据，传统的犯罪模型无法刻画三者间的相互作用关系，因此不能有效预测犯罪事件发生。随着GPS和智能手机等位置感知设备的普及，轨迹数据的获取已没有技术障碍。如何将轨迹大数据引入传统的犯罪模型是2016年将启动的重点基金项目的核心内容。期待通过该项目在犯罪模型研究上取得突破性的改进，进而为犯罪防控和案件侦破提供更有效的科学依据。



刘行健

(香港大学城市规划与设计系，博士，助理教授)

观点：新数据环境有助于对现有地理学理论的验证，大数据的生产、处理和使用受到其所处的社会自然、经济等地理环境的影响。

大数据中的地理(理论)：首先，大数据、开放数据与集成起来的小数据一起构成了地理学研究的新数据环境，这一新环境有助于对现有地理学理论的验证。例如，不少分析利用手机位置、网络流量与社交媒体签到等数据刻画城市空间结构与城镇体系，这些分析是对点—轴、增长极、中心地学说等传统理论在更细时空尺度上的验证。其次，新

数据环境将推动新地理学理论的发展^[6,7]。传统的地理学理论往往以均质化集合单元的长期变化作为分析对象，而新的数据环境下所涌现的即时、多维度、细粒度的观测将有助于关注异质化个体短期变化的地理学理论的产生^[8]。

地理（环境）中的大数据：一方面，大数据的生产、处理和使用受到其所处的社会经济自然等地理环境的影响，这一特性使得大数据对地理环境的表达是有选择性和片面的^[9]。例如，社交媒体数据（如微博）无法反映较少上网人群（如老年人）或社交媒体用户在特定时间段（如学龄儿童在校期间）的动态。类似的，由用户产生的大数据（志愿者地理信息，诸如OpenStreetMap、Geotagged Wikipedia以及带位置的微博等）较多记录于经济发达地区，而对乡镇或者欠发达地区的观测较少，造成大数据时代的“数据盲区”。另一方面，地理学研究也亟待探索日渐丰富的虚拟的数据环境如何影响并改变真实的地理格局与过程^[10]。例如，电子地图、电子商城与点评网站构建了现实世界在虚拟数字环境中的表达，每一次基于虚拟数据环境的出行、购物与消费行为都改变着现实世界中的人流、物流和信息流，促成数据对现实的间接改造。



刘瑜

（北京大学遥感与地理信息系统研究所，博士，教授；《Computers, Environment and Urban Systems》副主编）

观点：大数据的兴起为进一步定量地理解社会经济环境提供了一种新的手段；大数据存在数据代表性问题，作用不能夸大，依然需要去构建模型，解释大数据所蕴含的行为模式和地理规律。

传统的遥感技术利用光谱特征对地物进行观测，但无法有效地感知社会经济环境，而大数据所蕴含的社会经济特征，是对传统遥感数据的有力补充。最近提出“社会感知”概念及研究框架^[11]，指出社会感知就是指借助于各类大数据研究人类时空间行为模式，进而揭示社会经济现象的时空分布、联系及过程的理论和方法。大数据可从三方面感知人的时空间行为：①对地理环境的情感和认知（如基于社交媒体数据中可以获取人们对于一个场所的感受）；②在地理空间中的活动和移动（如基于出租车、签到等数据可以获取海量移动轨迹）；③个体之间的社交关系（如基于手机数据可以获取用户之间的通话联系信息）。社会感知的研究框架包括人、地、时三个基本要素。首先，在“人”的方面，社会感知数据可以获取人的活动与移动、社交关系、情感与认知等行为模式；其次，在“地”的方面，可以基于群体的行为特征揭示空间要素的分布格局、空间单元之间的交互以及场所情感与语义；最后，从“时”的视角，可以发现地理过程（尤其是人文地理过程，如城市空间结构演化）的规律和特征。社会感知由于其独特的“由人及地”的研究范式，为探究“人地关系”这一地理学基本命题提供了一个全新视角。在传统的人文地理学研究中，由于人的样本量通常较少，难以直接根据行为特征去推断地理环境。而根据社会感知框架，时空大数据使得基于人的时空行为模式感知并理解地理环境特征成为可能。回顾人文地理学发展的几次转向，计量革命因为缺少对人的关注而受到批评，而后期行为主义则受限于样本太小、缺乏定量模型。大数据的出现，在一定程度上弥补了上述两个方面的不足，对于地理学尤其是人文地理学研究具有重要支撑作用。

大数据也存在一些问题。首先，就目前地理研究中广泛应用的手机信令、交通轨迹和社交媒体数据等而言，大都存在数据代表性问题。例如，基于出租车数据，只能获取城市居民特定出行目的的信息，会过低估计通勤出行的比例。“大数据”并非是“全数

据”，在研究中，需要对于数据的代表性要有充分认识，处理大数据的代表性问题的一个可能的途径是集成多源数据交互印证。其次，大数据主要是提供了一种对既有模式及规律的观测手段，对于地理学的意义是毋庸置疑的，就如同显微镜的出现，使我们能够观察到以前看不见的事物（如细胞等），从而促进了生物学发展那样。但是，这个作用也不能夸大，依然需要去构建模型，解释大数据所蕴含的行为模式和地理规律。从时间维度上看，大数据主要是刻画了已经发生的过程，对于地理过程的实时描述以及对未来进行预测，同样需要与合适的模型进行结合。在这个过程中，基于访谈、调查问卷等途径获取的小数据的价值同样不可忽视。



龙瀛

(北京市城市规划设计研究院，博士，高级工程师；剑桥大学建筑学院马丁研究中心访问学者；《International Review for Spatial Planning and Sustainable Development》执行主编)

观点：大数据和开放数据所构成的新数据环境，将为人文地理学带来新的视角和发现；传统的城市研究地盘的概念，将会逐渐弱化，异地的学者也有望对千里之外的城市进行深入的人文地理研究。

中国人文地理学领域的研究，从空间尺度看主要有两类：一是针对单一城市较为深入的研究，例如研究广州城市贫困问题、北京的公共服务设施的配套水平；二是研究范围覆盖全国或多个省市地区的区域分析，研究单元较大，多以县域、市域或省域为单元。已有研究在研究尺度和粒度上往往难以做到两者兼顾，大的研究范围通常以牺牲粒度为代价，而高粒度的研究往往研究范围较小。

随着由大数据和开放数据所构成的新数据环境的形成，以及日益成熟的计算能力和日臻完善的区域、城市分析模拟方法，“大模型”成为城市和区域研究的新范式^[12,13]。大模型是由大规模数据驱动的定量城市与区域研究工具。利用简单直接的建模方式，兼顾了大尺度和精细化模拟单元，代表了一种新的研究范式。大模型方面，开展了若干实证研究，例如，中国所有城市的城镇建设用地范围识别、地块尺度的城市扩张模拟、城市增长边界评价、公交站点服务覆盖评价、人口对PM_{2.5}的暴露评价、功能性城市地域识别、城市次中心识别等，这些研究大多在全国范围开展，囊括了所有或大多数城市，研究力度多较为精细，如地块、街区、乡镇街道办事处等。

期待大模型这种自下而上的研究视角，且精细化、定量化、全面化的研究方法，为人文地理学带来新的视角和发现。可以预见，随着新数据环境的不断成熟，传统的城市内尺度的研究地盘的概念，将会逐渐弱化，学者有望对千里之外的城市进行深入研究。



陆锋

(中国科学院地理科学与资源研究所，博士，研究员；资源与环境信息系统国家重点实验室常务副主任；中国GIS协会理论与方法委员会主任)

观点：大数据很多具有地理空间本质，从big data到big geo-data的过渡很自然；大数据的本质不是超大的数据规模，而是异构性和动态性；地理学研究是大数据战略的天然试验场；基于地理大数据的地理计算与近期呈现的城市计算、社会计算交叉融合，直接应用于社会管理与普通大众的日常生活。

“大数据”并无什么特别之处，无非是以往较少被关注的信息，或者很难获取与处理的信息，在IT技术的推动下，进入量化分析平台，在不断细化的时空粒度下观测自然与社会环境。作为观察人类自身社会行为的“显微镜”和监测大自然的“仪表盘”，大数据无疑是地理学研究的福音，其价值不言而喻。人类生活中所产生的数据，有80%与空间位置有关，因此，很多大数据具有地理空间本质^[14]。从big data到big geo-data的过渡很自然。

物联网、云计算是基础设施的变革，而大数据是思维方式的变革。大数据的本质不是超大的数据规模，而是异构性和动态性，强调数据的不同来源、日积月累和群体参与。从这一点说，不存在“大数据”与“小数据”的界限，大小是相对的，也不存在跟风的担忧。地理学研究常用的遥感数据，数据规模海量增长，以往也从未冠之以“大数据”称号，只是到了移动互联网和物联网时代，“大数据”概念才深入人心，并迅速普及。

地理学是研究人地关系的学科，涉及自然环境与社会系统的方方面面。地理学研究是大数据战略的天然试验场。在大数据背景下，传感器网络、个体出行过程、网络行为、消费记录等均可能成为具有隐式地理空间形态的地理分析数据源，便于研究自然环境、社会动态、人口流动等^[15]。此外，传统地理信息注重的是地表要素的静态描述，其所附带的属性数据的时间变化并不频繁，而大数据时代的地理信息更强调地表要素的实时属性变化，特别是移动的个体或群体对象，其空间位置与属性的实时变化更为凸显，同时，数据时空粒度不断细化，从而使地理空间概念越发重要。

地理大数据的出现促进了地理计算（GeoComputation）与城市计算（urban computing）和社会计算（social computing）的交叉和融合。

很多城市问题和社会问题本质上就是“地理问题”。城市计算和社会计算都是涉及复杂系统、数据挖掘、网络科学、社会学、管理科学、自然语言处理、信息检索等多个学科的交叉研究领域，可以理解为大数据技术的城市与社会应用过程。其中很多研究场景和内容与大数据的空间隐喻息息相关^[16]。



秦承志

（中国科学院地理科学与资源研究所，博士，研究员）

观点：“大数据”被引入到地理学研究中是令人激动的，大数据+简单算法所呈现的效果，常常比以往“小数据”+复杂算法的效果更显著，但并没有上升到研究范式变化的高度。

GIS再一次从计算机科学、信息技术中获取“灵感”，“大数据”的概念被引入到地理学研究中，尤其是用于解决地理学中直接以人和社会为对象的各种问题，这是令人激动的。从概念理解的角度来看，目前有些地方将遥感数据也称为“大数据”，这是不合适的。海量数据并不意味着大数据，这种理解将大数据矮化为“新瓶装旧酒”。大数据和与其有关联的自发地理信息（volunteered geographic information, VGI）之间的差异值得琢磨。两个概念都以大量个体数据汇总后的增值应用为目的，VGI从个体数据产生之初，应用目的就是明确、统一的，或称应用目的先行；大数据的应用目的明确且多样，但个体数据的产生本身并不针对这些应用目的。这个特点使大数据具有更高的增值应用效率。

由于有了大数据，大数据+简单算法所呈现的效果，常常比以往“小数据”+复杂算法的效果更显著。是否上升到研究范式变化的高度上？这对于地理学研究意味着些什么？这些都还需要冷静思考。思考大数据能激发或推动地理学中问题的产生和解决，尤其是地理学研究者擅长的问题。否则，就只是GIS又一次从计算机科学、信息技术的借用，在激动之余不免留下更多的遗憾。



王慧

(厦门大学建筑与土木工程学院城市规划系，博士，教授；厦门大学海峡两岸城市规划研究所副所长)

观点：大数据对人文地理学的影响甚于对自然地理学的影响；目前大数据在地理学中的应用，尚未取得理论上的突破创新；期待大数据时代地理学新概念、新思想、新方法的诞生。

现代空间信息及测绘技术的发展，极大地改变了地理学认识地球、观察世界的视角与视野，由此造就了一个快速崛起的地理学分支——GIS，也使地理学成为最早开始大数据量级数据分析与应用的学科之一。

综合多维的国情地理数据，是一个大数据宝藏富矿，在优化区域布局、防灾减灾、城市精细化管理、智慧城市等方面发挥重大作用。地理科学急需经济社会发展导向、解决问题导向的综合国情地理数据开发应用，除了传统的综合对地观测与地理国情普查数据源外，百度、阿里巴巴、微博、Facebook、Twitter、手机GPS、车载GPS、社区监控摄像头等多源“城市传感器”正在日夜产生大量基于位置包含丰富个体“人”空间行为、社会行为细节特征的信息。这些大数据的充分挖掘分析，有望带来人文地理学研究内容与研究范式的巨大突破创新。相关研究为“以人为本”导向的城市规划、建设、管理等提供坚实的认知基础。大数据对人文地理学的影响甚于对自然地理学的影响。

目前大数据在地理学中的应用，还基本局限于校验印证既有理论，尚未取得理论上的突破创新。今后应该重视有关研究中思想和范式的提炼总结，尽早形成大数据时代地理学新概念、新思想、新方法的系统论述。



王鹏

(北京清华同衡规划设计研究院有限公司，高级规划师，技术创新中心副总工)

观点：大数据对地理学和城市规划学科的方法论的变革是颠覆性的；最明显的改变就是把学者们的精力从搜集和处理数据中解放出来；大数据包括行业传统数据的再认识和利用，尤其是现场调研、统计、测绘、遥感等数据，使用新的工具进行组织和结构化，可以使其具备大数据的特性和能力。

大数据带来地理学和城市规划学科方法论的变革，这个变革尤其体现在对人和社会的感知方式上。传统的城市空间研究方法主要是对物理空间的感知，而对人和空间社会的感知，往往是通过随机抽样方式，然后试图用模型去还原全貌，主要是研究精力很大程度集中在数据获取、抽样和还原方法上，而且往往独占性的数据成为学术研究的核心理资源。现在大量丰富的全样本行为数据的出现，最明显的改变就是把学者们的精力

从搜集和处理数据中解放出来，将精力放在提升用数据解决问题的能力，真正发挥数据的价值。

传统城市研究与规划的理论基础之一是统计学及线性预测模型，对于不确定性日益加强的复杂城市系统，新方法论将应对非线性、自组织的特征，大数据只是数据层面的一个改变，背后是一整套科学方法和工具体系，其代表就是机器学习、数据挖掘这些非精确的计算技术。从这个角度来看，与IT行业的智慧城市是殊途同归的。

大数据科研的一个致命问题是取决于拿到什么数据。就是单一的数据源对于描述问题过于片面和薄弱。三种以上不同类型大数据叠在一起的时候，会挖掘出非常有意思的事情，如人的移动、基础设施、经济运行。对于规划行业来说，优势是在原有的研究范式里，广泛接触和应用城市各部门的数据，具备综合使用数据解决问题的方法。但问题是，数据只是在规划编制过程中以碎片的方式出现，横向缺少协同，纵向缺少流动，需要行业全面解放思想，跳出传统规划、尤其是法定规划的窠臼，延伸和拓展研究对象和服务方式。



王伟

(中央财经大学政府管理学院，博士，副教授；中国城市规划学会城乡规划实施学术委员会委员)

观点：大数据兴起为“人地关系”研究提供了从“群体人共性”到“个体人个性”转变的可能；城市是以人为核心的复杂大系统，大数据为研究“以人为本”的城市打开了一扇窗户。

人地关系是地理学研究的核心，也是城市规划实践的核心。当前至少在发展理念上经历着从“土地城镇化”向“人本城镇化”的转型，但是现实的操作层面却仍然面临着“为人或为地”的博弈困境。“为人或为地”不单是规划手段的问题，而是关系到城市的基本功能取向。人和地之间是一种复杂的非线性关系，科学理解与平衡运用这种关系对于地理学与城市规划空间实践具有重要意义。而各种以“个体人”为粒度的大数据研究兴起无疑促进了从“群体人共性”到“个体人个性”的实现转变。然而以“人”为本，在大数据时代，面对两个挑战：

从研究的角度，以“个体的人”为本，则面临着城市整体利益与个体利益的平衡问题，移动互联网对未来社会的影响会进一步加深，整个社会呈现出原子化的碎化趋势。但如果过度分散，则城市所具备的集聚经济优势如何保持？从这个意义上讲，大数据应该是发现“千万个”人背后的“千万个个体行为的宏观规律”。从产业的角度，以“个体的人”为本，则意味着每个人都是各种生活所需品的消费终端，是各种个性化产品定制集成的源头，是企业利润的源泉。企业所需要的大数据应该是发现千万个个体“人”的个性化生活和消费特征的数据采集统计分析，来指导生产系统更加敏捷的响应。

这两大挑战，对地理空间会形成前所未有的冲击，意味着生活、生产、商业、物流、交通等都会发生重大变革。今日所面对的城市是以人为核心的更加巨大更加复杂的系统，大数据为我们打开了一扇窗户。然而在欣喜之余，还应多一份冷静：如何在大数据中平衡“工具理性与价值理性”？如何通过少数精英的“个体理性”推动更大社会“集体理性”的形成？



甄峰

(南京大学建筑与城市规划学院, 博士, 教授; 南京大学智慧城市研究院副院长; 中国地理学会城市地理专业委员会副主任委员)

观点: 城市的“大数据时代”已经到来; 大数据的挖掘及其应用已经成为国内外城市、地理与社会学研究的前沿; 城市地理与规划学者要主动强化数据意识, 激发大数据智慧。

数据是地理与城市研究的重要基础, 是分析城市发展现状、问题与特征的基本素材, 更是解释城市发展机制、科学规划城市空间增长的重要依据。作为人口和社会经济活动高度集聚的城市, 也是大数据产生和集中的场所。大量反映人类社会经济活动特征、人地关系作用过程的数据随时随地都在产生并逐渐被我们所感知、获取, 城市的“大数据时代”已经到来^[17]。

近年来, 大数据应用已经成为国内外地理学、城市科学与社会学研究的前沿方向之一, 涉及到人类社会经济活动的方方面面, 且研究理念、方法和技术也正在系统化^[18]。大数据对于城市研究与规划而言, 不仅仅是新数据源在城市社会经济领域的分析与应用, 更为重要的是, 它通过对城市复杂要素作用过程的分析和解读, 为推动城市研究与规划的科学化, 找到了一个可能突破的方向。

当然, 数据驱动的城市研究还处于实践摸索阶段。不能只强调时空大数据的表层信息挖掘, 还要探索深层数据以及对数据背后蕴含的人文特征进行深度挖掘。更需要强调的是, 大数据的城市研究, 还是要在适用的理论框架指导下, 通过对现象和特征的分析与讨论, 结合其他数据, 加强各类要素互动机制与规律研究。总之, 对于这样一个新领域, 城市地理与规划学者要主动强化数据意识, 激发大数据智慧, 跳出学科发展固有框架, 以应用研究需求为导向, 以(大小)数据整合与融合为动力, 推动城市理论与实践的创新发展。

主题总结

参与讨论的学者主要来自地理信息科学、人文地理、城市规划等领域, 并不能代表地理学领域对“大数据与地理学”的全面思考。欢迎更多对此主题感兴趣的学者同仁参加讨论。

通过参与笔谈学者的观点碰撞, 形成以下几点基本共识: 大数据不同于云计算、物联网、网格计算等传统IT领域的热潮, 它发端于互联网时代的信息技术, 但并不止于技术或数据, 它是一种思维方式和认知改造世界的能力提升; 地理学的发展离不开“互联网+”与大数据时代的影响与冲击, 它们给地理学的发展带来机遇与挑战; 大数据并不能改变和替代地理学的核心思想与基本范式, 地理学的核心命题是“地球表层系统”、“人地关系”与“时空变化”, 这决定了地理学研究是大数据应用与发展的天然试验场。

当然, 并不能奢望一次简单的笔谈讨论就能解决“地理学与大数据”碰撞所涉及的方方面面问题。参与讨论的学者的观点也针锋相对, 有的认为目前对“大数据”过于热炒, 有的认为大数据时代的地理学将发生深刻变化。还有很多问题没有答案, 如何将地理学的科学问题与基本原理与大数据方法、理论有机结合? 如何建立地理学大数据的数据体系与研究范式? 是否能够通过大数据驱动推进地理学与其他学科领域的交叉、促进

学科内部不同领域方向整合? 在技术层面、政策层面上如何改进地理学研究所需大数据获取与应用的方法途径? 相信随着大数据与地理研究案例的不断呈现, 理论思考与实践应用的逐步深入^[9], 各种问题都会逐渐得到答案, 而新的问题总会不断产生, 学科发展在这样的思考与实践碰撞中不断发展, 并服务于人类社会与自然系统。

最后, 借用鲁迅先生的名言——“其实地上本没有路, 走的人多了, 也就成了路”, 来结束本次笔谈。当新生事物出现时, 既不能盲目跟风, 也不能一棍子打死, 因为实践是检验真理的唯一标准。

致谢: 感谢《地理研究》编辑部朱晓华博士及城市数据派 (udparty.com) 对本次笔谈的大力支持。

参考文献(References)

- [1] National Research Council. Understanding the Changing Planet: Strategic Directions for the Geographical Sciences. Committee on Strategic Directions for the Geographical Sciences in the Next Decade. New York: National Research Council, 2010.
- [2] Sui D S. Introduction: Strategic directions for the geographical sciences in the next decade. *The Professional Geographer*, 2011, 63(3): 305-309.
- [3] 傅伯杰. 地理学综合研究的途径与方法: 格局与过程耦合. *地理学报*, 2014, 69(8): 1052-1059. [Fu Bojie. The integrated studies of geography: Coupling of patterns and processes. *Acta Geographica Sinica*, 2014, 69(8): 1052-1059.]
- [4] Collins J P. The fourth paradigm data-intensive scientific discovery. *Science*, 2010, 327(5972): 1455-1456.
- [5] Lazer D, Kennedy R, King G, et al. The parable of Google Flu: Traps in big data analysis. *Science*, 2014, 343(6176): 1203-1205.
- [6] Batty M. *The New Science of Cities*. Cambridge: The MIT Press, 2013.
- [7] Batty M. Big data, smart cities and city planning. *Dialogues in Human Geography*, 2013, 3(3): 274-279.
- [8] Graham M, Shelton T. Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography*, 2013, 3(3): 255-261.
- [9] Kitchin R. *Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage, 2014.
- [10] Rabari C, Storper M. The digital skin of cities: Urban theory and research in the age of the sensed and metered city, ubiquitous computing and big data. *Cambridge Journal of Regions Economy and Society*, 2015, 8(1): 27-42.
- [11] Liu Yu, Liu Xi, Gao Song, et al. Social sensing: A new approach to understanding our socio-economic environments. *Annals of the Association of American Geographers*, 2015, 105(3): 1-19.
- [12] 龙瀛, 吴康, 王江浩, 等. 大模型: 城市和区域研究的新范式. *城市规划学刊*, 2014, (6): 55-63. [Long Ying, Wu Kang, Wang Jianghao, et al. Big models: A novel paradigm for urban and regional studies. *Urban Planning Forum*, 2014, (6): 55-63.]
- [13] Long Y, Shen Z. Big models: From Beijing to the whole China. *Eprint Arxiv*, 2014, 116: 255-272.
- [14] 陆锋, 张恒才. 大数据与广义GIS. *武汉大学学报: 信息科学版*, 2014, 39(6): 645-654. [Lu Feng, Zhang Hengcai. Big data and generalized GIS. *Geomatics and Information Science of Wuhan University*, 2014, 39(6): 645-654.]
- [15] 陆锋, 刘康, 陈洁. 大数据时代的人类移动性研究. *地球信息科学学报*, 2014, 16(5): 665-672. [Lu Feng, Liu Kang, Chen Jie. Research on human mobility in big data era. *Journal of Geo-Information Science*, 2014, 16(5): 665-672.]

- [16] 余丽, 陆锋, 张恒才. 网络文本蕴涵地理信息抽取: 研究进展与展望. 地球信息科学学报, 2015, 17(2): 127-134. [Yu Li, Lu Feng, Zhang Hengcai. Extracting geographic information from web texts: Status and development. *Journal of Geo-Information Science*, 2015, 17(2): 127-134.]
- [17] 甄峰, 王波. “大数据”热潮下人文地理学研究的再思考. 地理研究, 2015, 34(5): 803-811. [Zhen Feng, Wang Bo. Rethinking human geography in the age of big data. *Geographical Research*, 2015, 34(5): 803-811.]
- [18] 杨振山, 龙瀛, Douay N. 大数据对人文—经济地理学研究的促进与局限. 地理科学进展, 2015, 34(4): 410-417. [Yang Zhenshan, Long Ying, Douay N. Opportunities and limitations of big data applications to human and economic geography: The state of the art. *Progress in Geography*, 2015, 34(4): 410-417.]
- [19] 隋殿志, 叶信岳, 甘甜. 开放式GIS在大数据时代的机遇与障碍. 地理科学进展, 2014, 33(6): 723-737. [Sui Dianzhi, Ye Xinyue, Gan Tian. Open GIS for big data: Opportunities and impediments. *Progress in Geography*, 2014, 33(6): 723-737.]

Geography interact with big data: Dialogue and reflection

WU Zhifeng¹, CHAI Yanwei², DANG Anrong³, GONG Jianhua⁴, GAO Song⁵, YUE Yang⁶,
LI Dong⁷, LIU Lin⁸, LIU Xingjian⁹, LIU Yu¹⁰, LONG Ying¹¹, LU Feng¹², QIN Chengzhi¹²,
WANG Hui¹³, WANG Peng¹⁴, WANG Wei¹⁵, ZHEN Feng¹⁶

(1. School of Geographical Sciences, Guangzhou University, Guangzhou 510006, China; 2. College of Urban and Environmental Sciences, Peking University, Beijing 100871, China; 3. School of Architecture, Tsinghua University, Beijing 100084, China; 4. Institute of Remote Sensing and Digital Earth, CAS, Beijing 100094, China; 5. Department of Geography, University of California, Santa Barbara, Santa Barbara 93106, USA; 6. College of Civil Engineering, Shenzhen University, Shenzhen 518060, Guangdong, China; 7. China Academy of Urban Planning and Design, Beijing 100044, China; 8. School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China; 9. Department of Urban Planning and Design, Hongkong University, Hongkong, China; 10. School of Earth and Space Sciences, Peking University, Beijing 100871, China; 11. Beijing Municipal Institute of City Planning and Design, Beijing 100045, China; 12. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China; 13. School of Architecture and Civil Engineering, Xiamen University, Xiamen 361005, Fujian, China; 14. Beijing Tsinghua University of the Same Scale Planning Design Institute Co., Ltd., Beijing 100085, China; 15. School of Government, Central University of Finance and Economics, Beijing 100081, China; 16. School of Architecture and Urban Planning, Nanjing University, Nanjing 210093, China)

Abstract: In the internet era, "Big data" wave spread rapidly to the economic and social fields. Geography is the natural laboratory in which big data research and application can be seen at work. The written speech focused on collision between geography and big data. It reviewed big data research and application in geography study. We also discussed the opportunities and challenges we would face during this collision. In summary, big data has had a certain influence on the geography research, especially in the human geography domain. Geographic information science will develop rapidly in the internet era of big data. But there are few disturbances in physical geography. Big data can not change the core proposition and the basic paradigm of geography. We should hold an open inclusive attitude to big data theory study and application research in geography.

Keywords: Geography; big data; opportunities and challenges; written speech; reflection