



在地理时空数据中寻找“未知”

——记国家自然科学基金创新研究群体项目“地理时空数据分析”

■本报记者 甘晓

从全国气象数据到空气质量数据、从高分卫星数据到海洋浮标数据、从出租车轨迹到室内人员定位数据、从导航软件中的拥堵信息到应用程序里的美食推荐，这些都是蕴藏丰富科学内涵的“地理时空数据”。

多年来，中国科学院地理科学与资源研究所资源与环境信息系统国家重点实验室的研究团队对地理时空数据进行深度挖掘分析，创建了一系列数据分析的先进方法，并用分析结果回答了城市规划、生态环境、公共健康等领域面临的现实问题。2015年，研究团队获得国家自然科学基金创新研究群体项目（以下简称创新群体项目）支持，项目于近期结题。

“大数据时代，我们面临‘未知的未知’，也就是，我们并不知道自己还有哪些未知的知识，而数据会告诉我们答案。”创新群体项目负责人、中国科学院院士周成虎告诉《中国科学报》。

超前布局闯前沿

早在新冠疫情暴发初期，如何为疫情防控贡献自己的力量是每一位群体成员思考的问题。“要想知道病毒传播路径、疫情扩散规律，在哪些地方传播风险高，本质就是探索特定时间、空间内人的聚集与移动的规律。”周成虎介绍，“这正是‘地理时空数据’的概念。”

当时，他们敏锐意识到，“地理时空数据”尤其是关于人的行为的数据有望成为疫情防控的重要支撑。于是，他们迅速“逆行”回到实验室，紧锣密鼓地开展攻关。团队通过梳理已有研究发现，个体轨迹和群体活动大数据能够对没有药物干预时的传染病扩散规律分析、风险评估方面发挥重要作用。

很快，2020年1月22日开始，他们组建了应急小组，在各地“健康码”出现之前就提出采用社交媒体及移动通信运营商数据提升防控能力与效率的建议，随着疫情发展，成功解析和模拟了疫情的时空趋势等关键内容，据此提出地理区域阻断与分级防控，深



团队骨干成员在实验室合影（左七为周成虎）。

甘晓摄

刻影响并显著提升我国疫情防控高科技水平。群体关于疫情防控的相关成果近日获得测绘科学技术奖特等奖（自然资源部下属）。

在创新群体项目成员看来，能够在疫情袭来的燃眉之急时迅速用上自己的研究成果，离不开整个团队长达20多年的超前部署。

这一步“先手棋”始于20世纪90年代后期，周成虎带领团队开始布局当时颇为前沿的数据挖掘研究。“互联网的浪潮刚刚袭来，我们认为，随着互联网发展和基于传感器的物联网发展，能够对物理世界采集到的数据越来越多，一定会迎来大数据的时代。”周成虎说。

研究团队曾从地震、降水等地理现象的数据入手，力图寻找不同数据反映事件的相关关系甚至因果关系。对此，让创新群体项目成员深有感触的是，科技创新要实现“四个面向”，基础研究应当具有前瞻性。

潜心打磨“锋利的刀”

从大数据分析中获得结果固然重

要，但在创新群体项目成员看来，面向未来的基础研究和应用研究，方法学创新更为关键。周成虎用一个生动的比喻阐述了方法的重要性：“如果你有一把‘锋利的刀’，当实际问题来的时候就用上了，用它解牛可以，改用它杀鸡也行。”

“时空地理流模式挖掘”就是团队多年来潜心打磨的“一把‘锋利的刀’”。地理流是指不同位置之间的物质、信息和能量的流动。例如，交通流就是一个标准的地理流，交通摄像头、手机导航软件等来源的海量交通大数据就是不同位置之间的地理流。

要从这样海量、复杂的数据中获得有效信息，难度堪比“大海捞针”。如何准确捞到想要的那根“针”，直接取决于“挖掘”方法——那把“刀”锋利不锋利。

“如何让流这种高维复杂的数据用统一、简洁的形式进行表达，是分析地理流数据的关键难题。”周成虎说。经过多年攻关，创新群体项目成员成功解决了这一难题。他们提出了流空间的定义，构建了流空间数学表达模型，并提出了流空间模式的系统性分类，同时构建了流的聚集尺度量化方法及流模式分解模型等。

这把“刀”在新冠疫情防控中，通过分析人流的特征显著提升了病例预测和风险分析精度。不仅如此，这把“刀”还被用于研究城市居民通勤模式、城市公园和医院可达性、城市居民出行模式等实际问题。研究结果表明，流空间以及流模式的分析不仅将成为地理信息科学新的分支，还有望为交通、规划、公共设施服务等决策提供理论工具。

此外，创新群体项目成员还取得了地表温度与蒸散发遥感反演与验证、连续曲面综合建模、异质空间统计等多项方法学创新。其中，“地表温度与蒸散发遥感反演及验证”获2019年度国家自然科学基金二等奖。

宽松环境育人才

多项原创性成果的取得，离不开研究团队的共同努力。

在该创新群体项目执行中，逐步扩充团队成员范围，形成3支队伍组成的人才梯队，包括中年队、青年队和研究生队。其中，中年队主要是核心骨干成员，青年队主要由40岁以下的年轻成员组成，研究生队则由研究生组成。

“一个篱笆三个桩”，周成虎表示，团队协作有效营造出利于安心科研的宽松环境。特别是面临考核压力的年轻科研人员，更需要有强大的团队支持，帮助他们度过瞄准一流科研成果长期攻关中的“困难时期”。

同时，在周成虎看来，基础科学对未知的探索有成功有失败，需要给年轻人提供允许失败的宽松环境。“群体不能失败，但个人可以。”他强调。

为此，该创新群体项目专门设立了6个“青年探索专题”，包括人群移动模式分析与位置服务推荐、基于大数据人口格网化方法与区域集疏分析、基于遥感的城市辐射环境与生态多样性、城市交通盲点分析、地理大数据与公共健康和地理大数据生态环境估计建模等，旨在由中年队带领青年队，使青年队静下心来钻研科学难题，

不惧失败。

在“青年探索专题”支持下，一批年轻人成长起来，“80后”群体成员唐荣林便是其中一员。他在地表蒸散发遥感反演的研究方向上取得进展，并获2019年国家自然科学基金优秀青年科学基金项目支持。

此外，在工作方法上，时间管理得到格外强调。长期养成记工作日记习惯的周成虎设计了一个《时间笔记本》发给团队成员，红色封皮上印有一句

箴言——“时间是人生中最宝贵的不可再生资源”。《中国科学报》记者在网页上看到，每日日程安排从凌晨4点开始，以半小时为单位。“学会管理时间是人生最大的学问。”周成虎指出。

“发展事业、报效祖国、成就自我、回报社会”，这是研究团队的16字指导方针。如今，创新群体项目成员正以实际行动践行，在地理时空数据中寻找“未知”，把数据分析结果“写”在祖国大地上。

《中国科学报》：创新群体项目成员如何看待基础研究与现实问题之间的关系？

周成虎：我认为基础研究一定和现实问题相关，当然这样的结论需要加上“时间”这个维度。也就是说，首先需要看清楚，要解决的现实问题是关系到今天的现实还是明天的现实。

有些基础研究可能与今天的现实无关，而有些就需要解决今天面临的现实问题，这类基础研究应当做到“知其所以然”。既要能够解决问题，又要通过理解问题本身，知道为什么可以解决，这样才能对解决其他现实问题有所帮助。例如，我们对地理流的研究就是“知其所以然”的基础研究工作，通过将现实世界中各种流进行抽象，得到地理流的数学表达和分析模型，然后才能像一把“锋利的刀”，用在各种行业中去解决现实问题。

《中国科学报》：创新群体项目成员如何看待理论与具体应用之间的关系？

周成虎：目前从基础研究到应用转化，周期越来越短。新冠疫苗的研发中，基础研究与产品开发的距离非常短，甚至可以看成是零距离。因此如何让理论研究的成果以最快速度在应用中发挥作用，推

进规模化的产业应用，是科技创新能力的重要标志。

经过长期的探索与实践，我们通过对数据尤其是遥感大数据的汇聚、分析加工、产品生成等，摸索出了从理论研究到产业化的新模式STEP，即科学（Science）—技术（Technology）—工程（Engineering）—产品与产业化（Product）多创新链的协同，推进从原始数据到数据产品再产生，并已在国土、环保、规划等行业领域取得了不少应用进展。

《中国科学报》：创新群体项目成员对这一领域的未来方向如何把握？

周成虎：基于时空大数据分析，我们正在与固体地球科学家合作，共同推动“深时数字地球国际大科学计划”。这项计划的科学问题是研究46亿年以来地球如何演化。解决这一宏大的科学问题，必须依靠大数据来分析。

这项计划中，我们提出一种新的科研范式“数据—知识—模式”，即从数据中得到知识，再形成知识获取的模式。未来，我们将进一步深化发展时空大数据的理论方法与技术系统，推动“数据—知识—模式”相结合的地理科学研究范式，变革地理科学。

看清脑联接 探索“小宇宙”

——记国家自然科学基金创新研究群体项目“生物医学光子学”

■本报记者 张双虎

“信仰是相信我们所未看见的，而对这种信仰的回报，是看见我们所相信的。”中国科学院院士、海南大学校长骆清铭有个宏大的梦想——看清大脑、弄懂大脑。

在国家自然科学基金的持续资助下，骆清铭团队凭着超凡的信念和执着，以“蚂蚁啃骨头”的韧劲，在20多年时间里，从冷门的交叉学科研究成长为颇具国际影响力的研究团队。近年来，他们的坚守和执着有了“回报”，团队建立的全脑定位系统（BPS），能够在亚细胞分辨率水平对全脑精细观测；他们的高清成像技术（HD-MOST），能清晰看到神经元的联接，实现了“在太阳旁看到星星”。

解构“最后疆域”

人类通过大脑认知世界，却对认知世界的大脑知之甚少。

人脑是一个拥有1000亿神经元和10¹⁵个神经联接构成的复杂神经网络。同时，大脑又是一块无法触碰的“禁地”，人们无法弄清里面有什么、到底怎样。因此，大脑被称为人类认识自然的“最后疆域”，脑科学研究也成为世界各国科技人员角逐的前沿。

“研究大脑的结构和功能，是人类理解自我存在、自我潜力与自我局限的重要过程。”骆清铭告诉《中国科学报》，“人们的呼吸、心跳、繁殖等基本功能，以及每天听到、看到、感受到外部的世界，无一不是大脑中数百亿神经元作用的结果。一旦大脑出现异常，将引发机体功能紊乱（严重时可危及生命）与精神性疾病（抑郁症、帕金森综合征等）。”

认识大脑，弄清这些神经网络是如何交织在一起的、在意识与行为产生时神经元“暗地里”进行了怎样的操作，科学家需要先看清大脑——绘制出脑联接图谱。

骆清铭解释说，绘制出脑联接图



生物医学光子学创新研究群体合影（前排左七为骆清铭）。

研究团队供图

谱可以帮助科学家在全脑范围构建神经网络，弄清楚神经元从哪儿来、到哪儿去，如何构成网络并行使什么功能，神经元有多少种类型，最基本功能单元是什么，神经网络有什么特点，如何从结构之“连接”到功能之“联接”。

“类似于芯片反向工程，基于脑联接图谱，通过对脑神经网络的提取与分析、整理，可以帮助科学家洞悉大脑工作原理、神经元及神经网络的的工作机制，助力类脑智能产品的设计。”骆清铭说，“从脑疾病防治角度看，有了脑联接图谱，‘维修’就方便多了。当然，我们还可以基于脑联接图谱来优化、增强、开发或应用脑的智能。”

在太阳旁看到星星

为“看清”大脑，全球科学家都在不懈努力。国际上关于脑图谱的研究有功能磁共振成像（fMRI）、电镜等。“fMRI无法分辨出每个神经元，像‘雾里看花’。而电镜只能看到神经元的局部，就像‘只见树木不见森林’。”骆清铭说。

2010年，骆清铭团队在多项科学基金和其他项目资助下，发明了显微

光学切片断层成像技术（MOST），通过对鼠脑3%~5%的神经元进行高尔基染色，看到了这些神经元间的联接情况，获得了世界上第一套小鼠全脑图谱。

“我们设定的目标是，花生米大小（约1立方厘米）的鼠脑，在全脑范围内每一处的体素分辨率都优于1立方微米（这样才看清每一根神经元）。”骆清铭说，“这需要把鼠脑切片1万多层，而切片薄到1微米是艰难的挑战。”

对鼠脑切片时，必须用自然界最坚硬的材料——金刚石。在项目研究初期，刀具损坏是个令研究人员头疼的问题。“只要一运行设备，刀刀很快就出现缺口，更不要说连续稳定切削上百个小时。”该群体成员李安安说，“造成刀具损坏的原因有很多，如刀具、切削样本、切削参数、切削系统刚性、外部振动影响等。由于此前没人做过这件事，所以所有可能的原因都需要我们去一一排除。”

而且，这种金刚石刀当时只能从国外定制，往往是设计图寄过去，经历数月才能拿到刀具，这给研究带来很大的困扰。

在长期探索中，团队不断修正技

术路线，提出新的实现方式。2013年，他们建立了荧光显微光学切片断层成像方法和技术（fMOST）；2016年，实现双色成像（dmOST），既看到神经元也能看到神经元所在的位置，建立了一种类似全球定位系统的全脑定位系统（BPS）；2021年，他们又提高图像信噪比，就像能“在太阳旁看到星星”一样，发展出更高清的成像技术（HD-MOST），更清晰地看到神经元的联接，并解决了庞大数据量压缩等系列难题。

“该研究最大亮点是全脑定位系统技术。”骆清铭说，“这是目前在亚细胞分辨率、数据质量最好的成像技术，并在国际上保持了十余年的领先优势。”

学生的学生已成老师

在该群体宽大的实验室两侧，分别挂着中国科学院院士杨福家和杨叔子的

题词——“追求卓越”“为国争光”。

“我们也把这（题词）作为宗旨。”该群体成员李向宁说，“研究中遇到的困难和问题非常多，只有统一思想、提高认识，才能心无旁骛专心攻关。”

高端科研仪器研制技术门槛高、工程要求复杂、创新难度较大，很多技术难点需要三五年甚至更长时间才能突破，标志性成果更是滞后多年才能发表。群体成员中，有老师因此职称迟迟得不到晋升，也有研究生因此延迟毕业。但他们没人计较个人得失，以“蚂蚁啃骨头”的韧劲，克服一个个困难，最终实现团队的目标。

“全脑定位系统的每个样本在数据采集过程中实际需要完成百万次显微成像。”该群体成员袁菁说，“研发过程中，始终要求每一个细节都向100%逼近，时时思考是不是最优方案，任何环节的放松都可能影响最终性能，这个过程非常折磨人。”

《中国科学报》：目前我国在脑科学研究领域处于什么水平？

骆清铭：总体来讲，我国脑科学研究与发达国家尚有距离，但在个别方向上，我们已经和国际先进水平并跑甚至领跑。

《中国科学报》：您的团队目前是否开始研究人类大脑，弄懂大脑对人类意味着什么？

骆清铭：脑科学是人类科学的最前沿。目前科学家还没有正式开始在介观水平研究人类大脑。

弄懂大脑，意味着我们能看清大脑的宏观、介观与微观结构，解读神经元的联接及信息交流过程，理解多种尺度下神经网络动态行为，以及揭示大脑的信息处理机制等，甚至意味着有望建立一套完善的理

论框架来解释人类智能的形成。

弄懂大脑将揭开思维和意识的神秘面纱，会对人类的健康、教育、科技、社会发展等方面产生重大影响。健康方面，将有利于进一步促进神经性和精神性疾病分类；有利于探索诊断、治疗、治愈甚至预防阿尔茨海默症、帕金森综合征等疾病的有效方法；将为具有学习、睡眠、情绪、智力等障碍的人士提供精准医疗、服务及个性化的干预手段等。教育方面，将为推进人脑认知功能的开发和保护提供科学理论基础，有利于制定个性化的教育方案与评测机制，进一步提高人口素质。科技方面，将推动类脑计算、类脑芯片及脑机智能技术的发展，有利于模拟脑功能并发展更多人工神经网络算

法，实现更高级的人工智能；促进智能医疗器械、虚拟现实和可穿戴等技术的发展，带动相关研究领域及产业的创新发展，开启新一轮工业革命的新征程。社会方面，有助于回答生命和智力的起源问题，进一步为人类相关的应用开辟更广阔的前景，更好地满足人民对美好生活的向往。

《中国科学报》：请谈谈脑科学的发展方向，团队下一步的研究重点是什么？

骆清铭：工欲善其事，必先利其器。生命科学的每一次发展都以重大技术的进步为前提，脑科学研究也不例外。我们会继续聚焦脑科学研究的重大需求，以创新技术的发展为牵引，力争取得更多原创性成果。