中国科学院学部
科学与技术前沿论坛简报
第 130 次

学部工作局科学普及与学术处
《中国科学》杂志社 编报

2023 年 2 月 16 日

“数据驱动的新地学”论坛综述

一、国内外发展现状

大数据是知识经济时代的战略高地，是人类的新型战略资源。作为大数据重要组成部分的地球大数据，正成为地球科学的一个新的领域前沿，为推动地球科学的深度发展以及重大科学现象的认知发现正在作出贡献。

近年来，国际上推出了一系列与地球大数据相关的研究计划。美国的“地球立方体”项目，寻求以整体视角审视地球系统并创建管理地球科学知识的基础设施；欧盟启动的“活地球模拟器”研究项目，依赖于开放的标准，为所有地球科学领域提供多维数据和可扩展的服务；澳大利亚在启动的数据立方体项目基础上，最近又推出了澳大利亚“数字地球”项目；2017 年，俄罗斯正式启动“数字地球”计划，包括发射一系列的地球观测卫星，为该计划提供源源不断的数据源。

中国科学院充分认识到地球大数据的重要性，2018 年初设立了战略性先导科技专项（A 类）“地球大数据科学工程”（CASEarth），系统开展地球大数据研究。CASEarth 的宗旨即利用地球大数据驱动跨学

科、跨尺度宏观科学发现，以系统性和整体性的理念去研究一系列重大科学问题，在对地球系统科学认知上有重大突破，同时在决策支持上实现新的跨越，持续在科学发现、宏观决策、技术创新和知识传播等方面产出重大成果。2019 年由中国科学家倡议、13 个国际组织与机构共同发起的国际大科学计划——“深时数字地球”（DDE），致力于搭建全球地球科学家与数据科学家合作交流的国际平台，推动地球科学在大数据时代的创新发展，促进深度数据驱动的地球科学研究范式，造福科学、社会和可持续发展。

相比于欧美等发达国家，我国在地球大数据的研究中更偏重于数据集成、信息挖掘和大数据应用分析的平台建设。目前，包括中国科学院、北京师范大学、中国地质大学（北京）、南京大学等多个科研团队在大数据中心、大数据信息挖掘、大数据云平台等方面的研究都已经取得重要突破。

二、论坛概况

2022 年 10 月 29～30 日，中国科学院学部“数据驱动的新地学”科学与技术前沿论坛在中国科学院学术会堂召开。本次论坛由中国科学院学部主办，中国科学院学部学术与出版工作委员会、中国科学院地学部、中国科学院信息技术科学部承办，《中国科学》杂志社及青藏高原地球系统与资源环境全国重点实验室协办。论坛由陈发虎院士和徐宗本院士担任执行主席，邀请了 13 位中国科学院院士，以及来自中国科学院、西安交通大学、中国地质科学院地质研究所、清华大学、南京大学、中国地质大学（北京）、北京大学、南京师范大学、兰州大学、中国科学技术大学、武汉大学等高校和科研院所的 50 余位专家共同参与研讨，吸引现场和线上参会专家学者总计 400 余人。

本期论坛是一期典型的跨学科交叉论坛，旨在通过研讨和交流，集思广益，进一步明晰地球系统科学的研究方向，推动我国在大数据时代的地学研究变革中占据引领位置。

三、论坛召开的目的、重要性和必要性

随着科学技术的飞速发展和社会需求的强大驱动，人类产生的数据量正在呈指数级增长，逐步迈进了大数据时代。作为新型战略资源，大数据已成为知识经济时代的战略高地，推动了科学研究向数据密集型科学范式的转换，逐步蜕变为科学发现的新引擎。作为大数据的重要组成部分，地球大数据具有大数据的一般性质，同时具有很强的时空关联和物理关联，是当前地学研究的前沿领域。在学科发展和科学发现上，地球大数据带来了方法论和研究视角的创新，催生了地学方法论和思维模式的革命性变化，为地学研究提供了重要的发展机遇。另一方面，得益于深度学习、大数据分析、高性能计算的发展，以机器学习为核心的人工智能算法研究和应用迎来新的高峰。在此背景下，人工智能领域的技术进步为地学相关领域研究的智能化发展和融合创新带来了新的机遇和挑战。

目前,地学发展日益呈现密集数据驱动和多学科深度交叉的趋势。一方面，新技术和新方法的出现和应用推动了质量更高的地学数据呈指数级增加，另一方面，地球大数据和人工智能技术的发展推动了新研究范式的产生。如何更好地利用新技术获取高质量的地学新数据，如何更有效地构建大型地学数据库，如何更好地利用人工智能技术推动新地学的发展？探讨这些问题对于我国乃至全球地学的发展，都具有极其重要的意义。希望本次论坛在地球大数据的背景下，剖析大数据服务地学科学发现和重大科学计划的顶层设计和方法论，进一步甄别地学研究的优先和跨学科研究领域。同时，通过与人工智能的交叉融合，旨在提升对地球系统动态感知、智能推理和知识发现能力，进而探索新时期地学研究的前进方向。

四、报告内容

本次论坛围绕“数据驱动的新地学”这一主题，深入研究地球大数据与人工智能的深度耦合，探索数据驱动下的地球系统科学变革。

着重围绕“大数据、机器学习与地学”、“数据驱动下的地球系统科学变革”等交叉学科领域组织专题的评述报告，并进行充分研讨。希望通过“数据驱动的新地学”论坛这一平台，集思广益，进一步明晰地球系统科学的研究方向，推动我国在大数据时代的地学研究变革中占据引领位置。

议题一：大数据、机器学习与地学

中国科学院院士徐宗本作了题为“机器学习：从先验假设到自动化”的报告。他指出，机器学习是人工智能的最基础、最核心技术（算法），但机器学习的执行通常都是以一组基本的先验假设为前提的。报告分析了这些假设的作用、局限及其影响，提出突破这些基本假设的途径与方法，并举例说明新突破带来新价值。机器学习的当下应用还是“人工”的：人工标注数据、选择数据，人工搭建网络、选择算法，人工切换任务、适应环境；仍处于“选择”阶段，还没有达到“自主设计”、“创新设计”水平。随着人工智能技术的发展，机器学习必须要从人工化，走向自动化，迈向自主化。在这样的AI发展大趋势下，让机器学会人类的学习方法论，或者更严格地说，学会模拟学习方法论（Simulate Learning Methodology，SLeM）成为一种必需和趋势。报告还严格定义学习方法论问题，提出SLeM的双层优化模型和“超参数化”求解方法，建立SLeM的基本理论，并展示对多个机器学习自动化问题的成功应用。

中国科学院院士陈发虎作了题为“青藏高原地球系统观测、数据与模拟”的报告。他在报告中指出，青藏高原是我国重要的水资源安全屏障、生态安全屏障、国土安全屏障、战略资源储备基地、中华民族特色文化保护地。青藏高原以其强烈的多圈层相互作用闻名于世，是国际地球系统科学研究的天然实验室。研究青藏高原的形成、演化与影响，是当前国际地球系统科学研究的前沿领域。青藏高原碰撞隆升导致地形地貌、大气环流、自然环境格局和生物与人类演化产生链

式响应。同时，青藏高原是受全球变化影响最剧烈的区域，其变化进一步影响着区域乃至全球的气候环境变化。青藏高原地球科学研究取得了丰硕成果，但从地球系统科学研究角度看，还存在瓶颈：（1）动力演化和现代气候变化过程研究之间时间不联通；（2）六大圈层间的关键动力过程和相互作用机制缺少深入研究；（3）需要立足青藏高原，联系南北极，在全球视角下形成地球系统科学理论。因此，新时期青藏高原地球系统科学研究需要打通时间隧道、圈层隔离和空间局限，聚焦高原生长、全球变化、人类活动三重驱动下的多圈层链式响应及影响研究。面向青藏高原地球系统科学研究和可持续发展对观测平台、数据平台和模型平台的需求，亟需建立青藏高原地球系统观测-数据-模型一体化平台：（1）分阶段建立青藏高原国内与国际观测联盟，构建青藏高原天-空-地一体化观测平台；（2）以国家青藏高原科学数据中心为信息枢纽，实现观测数据的实时汇聚、实时融合，服务模型发展和决策支持；（3）分阶段发展青藏高原气候系统模型、地球表层系统模型和地球系统模型，助力青藏高原地球系统理论突破，形成针对青藏高原可持续发展复杂问题的决策支持能力。

中国科学院院士张平文作了题为“机理与数据的融合计算”的报告。他通过剖析应用数学（含统计学）的演进过程，强调融合计算已经成为当前科研和创新的重要范式；介绍了融合计算的若干研究模式，及其为应用数学发展所带来的增长点；通过精准气象预报和分子动力学模拟两个案例来详细阐述机理与数据融合计算的实践过程；最后对融合计算的未来发展趋势进行探讨。

议题二：数据驱动下的地球系统科学变革

中国科学院院士郭华东作了题为“地球大数据思考”的报告。他在报告中指出，科学大数据具有数据密集型范式的特点，具有数据的不可重复性、数据的高度不确定性、数据的高维特性、数据分析的高度计算复杂性等内部特征。利用大量数据的相关性可取代因果关系和

理论与模型，基于数据间的相关性能够获得新知识、新发现。作为少量依赖因果关系，而主要依靠相关性发现新知识的新模式，大数据已成为继经验、理论和计算模式之后的数据密集型科学范式的典型代表，带来科研方法论的创新。作为科学大数据重要组成部分的地球大数据，正在成为地球科学和信息科学交叉的新兴前沿研究领域，可为推动地球科学深度发展并产出重大科学发现作出贡献。地球大数据是一种典型的科学大数据，它由与地球有关的数据构成，包括地球的陆地、海洋、大气及与人类活动相关的数据。地球大数据是具有空间属性的地球科学领域大数据，主要产生于航天与地面大型科学实验装置、探测设备、传感器、社会经济观测以及计算机模拟过程。它一方面具有海量、多源、多时相、异构、多尺度、非平稳等大数据的一般性质，同时具有很强的时空关联和物理关联，具有数据生成方法和来源的可控性。地球大数据正成为认识地球的新钥匙和知识发现的新引擎。可为地球科学深入研究带来不竭的动力，可期待地球大数据驱动学科的创新与持续发展，可服务区域与全球可持续发展、气候变化、粮食安全、生态环境、风险评估和灾害应对以及为文化自然遗产保护与发展等提供科学决策支持。

中国科学院院士沈树忠作了题为“重建地球生物多样性历史的思考与方案”的报告。他在报告中指出，地球自现代生态系统建立至今已有5亿多年的演化历史，生物经历了由简单到复杂、由少到多的演化过程，什么因素决定了物种多样性？重大环境和气候变化事件如何影响生物多样性？回答这些重要科学问题，首先需要建立地质历史中海陆生物多样性的精细模式。地球上曾经生活过的生物99%以上已经灭绝，深时生物多样性重建依赖于不完整的化石记录。然而，受生物保存差异、岩相变化、采集程度影响较大，如何通过不完整的化石记录重建地质历史时期生物多样性的演变历史是一个重大科学难题。目前国际主流的方法是基于已发表的文献建立数据库,划分统计时间段，

进行全球或区域大尺度的多样性统计。但受限于方法自身的局限，此类研究的时间精度低（分辨率通常在千万年左右），因此，难以揭示重大生物事件的精细过程，无法识别突发性生物多样性灾难事件。最新的研究表明，深时生物多样性研究的时间分辨率必须至少达到十万年级别以下才有可能真实地揭示生物多样性演变中的重大突变事件，并使得与近代地球生物多样性变化作比较研究成为可能。基于高精度综合地质时间轴和同位素地质年代学的定量地层学和人工智能算法是解决这一问题的重要途径。

中国科学院院士王成善作了题为“深时数字地球：国际大科学计划回顾与展望”的报告。他指出，地球是复杂的四维系统，既包括现代三维空间系统（如深空、深海和深地），同时也包括 46 亿年地球演化系统（深时）。站在破解地球和生命演化的科学角度，站在数字革命的时代角度，站在中国应该为全球做出更多贡献的全球角度，我国科学家牵头发起了“深时数字地球”国际大科学计划（Deep-time Digital Earth，DDE），并获得了国际社会的积极支持和参与。DDE 计划的使命是整合地球演化全球数据、共享全球地学知识，愿景是推动地球科学研究范式的变革。针对生命、物质、气候和地理演化等地球演化领域的重大科学问题，DDE 计划依托我国在通信网络、高性能计算等方面的优势，通过整合数十亿年地球时空大数据，打造我国地球科学领域内首个自主研发、自主可控的全球数字基础云设施，实现地球演化创新突破，为全球地学界提供社会公共产品。DDE 计划的理念、模式被国际接受，形成多层次参与机制，组织架构与机制设置日趋完善。成为国际地质科学联合会设立大科学计划后认可的首个大科学计划，创始成员 24 家，包括英国地质调查局、俄罗斯地质调查局、印度地质调查局等国家行业管理部门和美国石油地质学家学会、国际地层委员会、国际数学地球科学学会等国际学术组织。设立由全球顶尖科学家和科技管理专家组成的管理委员会、执行委员会和科学委员会。DDE

计划利用“大知识、大数据、大平台、大科学、大传播”开展数字化时代的科学研究，在“开放科学”时代用云计算的思路打造数字化基础科学研究设施，用计算能力赋能地球科学家以打造全球开放环境、缩短科研创新周期、促进学科实现研究范式变革。2022 年，DDE 计划同联合国教科文组织、联合国基础科学促进可持续发展国际年、国际地质科学联合会共同院士团队在巴黎联合国教科文组织总部举行了 DDE 开放科学专题论坛（DDE Open Science Forum），得到国际科学界和工程界的积极响应。

中国科学院院士成秋明作了题为“‘数据-数学-超算’三位一体驱动现代地球科学创新”的报告。他在报告中指出，大数据、互联网、云环境和人工智能等革命性数字技术的快速发展和应用，使地球科学家能够研究地球如何随时间演化，重大地质事件如何发生和改变地球，以及与人类生活和社会发展密切相关的资源、能源和地质灾害分布情况。使用各种现代天-空-地-海技术获取的地球大数据代表了人类时间尺度上的当前地球，但从记录地球地质历史的岩石样本中分析的“长尾”数据往往以无格式文本、表格或地图的形式分散在文献中。机器阅读、机器学习和知识图谱技术可以在提取和整合这些类型的长尾数据以形成表征地球随时间变化的深时序列方面发挥重要作用。然而，这些类型数据的质量会受到多种因素的影响，包括但不限于岩石样品的不完善以及地质构造运动可能造成的岩石损失，因此对深时数据的分析需要使用适当的数学模型和方法及其软件进行特殊处理。局部奇异性分析（LSA）是最近开发的基于分形密度新概念的方法。LSA 可用作高通滤波技术，通过去除深时数据的趋势来有效提取弱异常。应用 LSA 分析全球同位素年龄数据、沉积物通量数据以及矿物和矿床数据的几个研究案例揭示了与板块构造演化和超大陆形成相关联的岩浆通量、沉积通量以及大气氧气等演化的幂式性、聚集性和周期性。证明大数据、数学和超算的集成可为深时数字地球研究提供强大的工具。

报告还简要介绍了深时数字地球（DDE）大科学计划，这是国际地质科学联合会（IUGS）为促进数据驱动地球科学发展而发起的一项新计划。

中国科学院古脊椎动物与古人类研究所研究员倪喜军作了题为“大数据背景下的人类起源模型”的报告。报告指出，狭义的人类起源是指人猿分异之后人类支系的演化，人类支系的演化方向，表现为获得直立行走的能力、脑容量增大、大规模制作和使用工具、发展复杂社会、形成文化传承，直至发展成为现代人类。广义的人类起源包括灵长类的起源、类人猿的演化、大猿的分化和人猿分异。自灵长类起源以来的演化过程，形成了人类灵活的手脚、彩色视觉、脑容量增大、社群结构复杂等人类的生物学特征。涉及人类起源的数据包括现生与化石灵长类和人类的解剖学特征、行为学特征、地质年代和地理分布、现生灵长类和人类的基因组、蛋白组以及少数化石种类的不完整的遗传信息片段。这些数据的特征在于化石记录不完整、保存信息有限、演化时间长、地质年代误差范围大、数据来源差异巨大、待检验模型的数量是天文数字等方面。目前用于模型选择的标准包括遗传距离、简约性、最大似然和贝叶斯推断。人类起源研究中目前面临的具体挑战包括化石准确定年、形态特征矩阵建立工作量巨大、异质性数据整合困难。在形态数据获得方面，尽管各种成像手段已经可以获得海量数据，但是提取解剖特征主要还是依赖于研究者的主观经验。在异质性数据整合方面，主要在于来自于分子生物学的遗传信息和传统的形态学信息属于不同系统层面的信息,两者目前缺少实质性关联。未来的工作应该针对这些具体挑战开展世界范围内的灵长类和人类化石地点的定年工作，开展标准化的形态特征数据采集，通过广泛的国际合作建立数据平台，开发新的数据分析方法，检验更多的演化模型。

中国科学院青藏高原研究所研究员李新作了题为“地学知识约束的机器学习与大数据同化”的报告。他在报告中指出，科学大数据推

动了科学研究向数据密集型科学范式的转换。针对地球系统科学研究中涉及的几种前沿大数据分析方法，如深度学习、物理约束的机器学习、因果推理和深度强化学习等，探讨这些方法的基本原理及其在地球系统研究中的应用前景和面临的挑战。其中，深度学习通过多层神经网络从数据中挖掘更深、更抽象的特征，极大地增强了模型的学习能力，在地球系统科学分类、回归、预测和预报等相关领域取得了巨大成功。物理约束的机器学习通过融合机理模型和地学知识，提升了机器学习模型的可解释性和移植能力，极大地增强了机器学习在地学知识发现和决策的有效性。因果推理侧重于从地球大数据中寻找地球系统变量间的因果关联，识别因果路径，量化因果效应，能够在地球系统的过程理解和反事实预测方面发挥重要作用。深度强化学习将深度学习的感知能力和强化学习的决策能力相结合，能够在高维环境中进行更强大的特征表达和非线性建模，在应对地球系统科学的决策问题中发挥重要作用。数据同化是增强地球系统可观测性和预测性的重要手段，已经广泛用于地球系统科学研究的几乎所有分支领域。随着地球大数据和大数据分析的快速发展，作为一种新的同化方案，大数据同化旨在将超高分辨率的地球系统泛在观测（包括人类系统的社会感知等非主流观测数据）融入超高分辨率地球系统模型，获取系统和参数的最优估计，提升地球系统模型的预测能力，为构建地球数字孪生提供重要的模型基础。展望未来，大数据分析方法有望为大数据驱动的地球系统科学研究提供关键手段，但机遇与挑战并存。

武汉大学资源与环境科学学院教授沈焕锋作了题为“地学机理模型与机器学习模型的耦合范式”的报告。他指出，构建物理驱动的机理模型一直是地学研究的核心科学范式，但仍然存在机理认知局限、欠定系统、计算负担等瓶颈问题；发展数据驱动的机器学习模型是地学研究范式转换的重要途径，但泛化性、迁移性、可解释性不足是其主要问题。耦合机理模型与学习模型则可以实现“理性主义”与“经

验主义”的结合，是当前最受关注的研究热点之一。然而，将显式的机理模型与隐式的学习模型进行耦合存在诸多挑战，尽管当前已经取得了一些研究进展，但仍然缺乏标准、统一的范式框架。该报告针对遥感反演与动力学数值模拟，分析机理模型与学习模型的内在瓶颈及其互补性，阐述以机理级联学习、学习嵌入机理、机理融进学习为核心的耦合范式框架。机理级联学习是将两种模型进行前后串联，一种模型的输出作为另一种模型的输入；学习嵌入机理，是以机理模型为主、学习模型为辅，将学习模型嵌入到机理模型之中，对原有的不确定过程进行替代或优化；机理融进学习，就是以学习模型为主框架，将物理知识融入其中，从而实现对学习过程的约束引导。报告在对耦合范式框架进行系统分析的基础上，进一步以地表温度、大气 $PM_{2.5}$ 等参量估算为例，介绍提出的时空关联学习、地理加权学习等方法，以及对地观测、模型模拟、社会感知等数据的融合应用方法。机理模型与学习模型的耦合，可以有效调整机理模型的“偏见”，避免大数据学习模型的“傲慢”，有效提升对地球表层参量的估算精度与效率，在地球系统科学研究、资源环境问题应对中展现重要的支撑能力。

西安交通大学研究员毕海霞作了题为“基于 AI 的遥感图像分类与识别”的报告。她在报告中指出，遥感图像分类与识别是遥感对地观测的关键任务。近年来，世界各国成功发射了大量对地观测遥感卫星，获取了海量遥感图像，遥感大数据时代已然来临。但是，目前遥感技术应用的瓶颈在于其图像处理技术远远滞后于其图像获取技术。因此，如何对遥感图像进行快速准确地解译，具有非常重要的理论和实际应用价值。近年来，蓬勃发展的人工智能技术将遥感图像分类与识别推向自动化、智能化。然而，基于人工智能的遥感图像分类识别任务面临两个巨大挑战，即：（1）小样本：尽管存在海量数据，但数据标记成本高，标记数据难以获取；（2）数据质量差：遥感图像中存在噪声大、样本不均衡及数据缺失等问题；（3）多源异构：多模态数

据的多源异构性为数据融合带来了挑战。为解决小样本问题，如何充分利用无标记数据，并在分类过程中有效融入遥感机理及领域先验为亟需解决的科学问题；为解决数据及标签噪声导致的分类精度差的问题，对数据噪声进行建模及对标签噪声进行检测及矫正为有效的解决思路；如何对模态间的关联性机制进行建模是解决多源异构数据融合的关键问题。该报告围绕以上研究课题介绍遥感图像分类识别中的代表性工作；同时对遥感图像分类识别的未来发展趋势进行展望。

五、论坛达成的共识及同领域学科发展建议

本次论坛报告精彩、讨论热烈。会议执行主席陈发虎院士总结了研讨中达成的共识，明确指出关注数据驱动的新地学研究范式必须将地球大数据和大数据分析方法（技术）并重，一方面收集、整合、管理好地球大数据；一方面创新地发展和使用大数据分析方法，深度挖掘数据背后隐藏的地学知识。这就需要我们践行“全链条”管理理念，从数据的采集、管理、共享、分析、应用到效益分析，全方位做好数据支撑和数据服务，做好数据中心基础建设。同时，从数据应用层面，需要深度整合数据资源，提高基于数据的信息挖掘和知识发现能力，真正落实地球大数据对地学学科发展的支撑和引领作用。

在政策引领方面，建议国家牵头组织开展交叉性科学研究，推出比较明确的研究所、高校、企业联合的科研项目，加大地球大数据相关的专项经费支持。近几年，虽然国家自然科学基金委成立了交叉领域学部，部分企业也参与到对交叉领域的经费支持中，但涉及“数据驱动的新地学”交叉领域的项目仍然有限，支持原始创新的经费偏少。目前，数据驱动的地学研究项目多偏向应用型研究，包括机器学习、深度学习等方法在地学领域的应用探索，在数据的生产、整合、管理、高效应用和知识挖掘方法方面的支持依然不足。

在科学研究方面，持续推进学科和综合性数据中心及云平台建设，开发、整理、完善系统的数据分析工具集，从理论到实践，全方位提

升数据对地学学科的支撑能力。根据文献调研和论坛讨论，目前对地球大数据的应用依然处于以相关性研究为核心的初级阶段，部分研究者偏向机器学习和深度学习等方法的初步应用，不重视理论层面的基础创新，应该制定相关的发展纲要，支撑基础分析理论的创新发展并长久实施下去，推进地球大数据与大数据分析理论的有机结合，实现数据支撑学科理论发展的宏伟目标。

在人才培养和平台建设方面，我国在机器学习/深度学习领域中各类专科人才不断涌现，但交叉领域特别是地球大数据与大数据分析理论交叉领域的人才队伍、团队建设和合作共享资源相对稀缺，缺少针对交叉领域人才的培养机制和相应的政策保障。建议国家投资建设有利于自主研发的创新平台，制定符合我国国情的交叉领域人才培养机制，真正地推进地球大数据对地学研究的实际支撑。

六、与国外同领域研究的比较

尽管在数据驱动新地学前沿交叉领域的一些前沿热点问题上我国的科学研究成绩突出，但从整体上来看，我国在该领域的引领发现偏少，在国际学术界地位有待加强。为了实现追赶和超越，我国需要充分发挥自身在地球大数据和大数据分析领域的研究优势，有序加强地球大数据专家、大数据分析专家之间的有效合作，推动交叉领域研究平台建设，将基础大数据和大数据分析技术结合，争取在数据驱动的新地学领域取得引领性优势。

我们要综合国际发展导向，落脚我国发展实际，制定符合我国国情的管理政策及发展方针。首先，我国基础大数据起步较晚，在卫星遥感观测、遥感算法反演领域距离国际先进水平还有差距，需要脚踏实地，认真落实基础理论研究，抓住大数据“风口”，完成大数据基础建设。其次，我国的大数据分析方法更关注于应用层面，对基础理论和算法研发的关注不够，多数时候依然处于追赶状态。建议从过程层面进行政策倾向，争取迎头赶上。最后，交叉学科建设是实现数据驱

动的新地学的重要途径，必须因地制宜，着眼实际，制定出具有我国特色的交叉学科发展方案，推动数据与技术的深层融合，助力新地学的学科建设。

（作者：陈发虎，中国科学院院士，中国科学院青藏高原研究所研究员；徐宗本，中国科学院院士，西安交通大学教授；李新，中国科学院青藏高原研究所研究员）

联系方式：中国科学院学部工作局科学普及与学术处，010-59358366