

科学数据，走向独立

中国科学院数据期刊集群发布

■本报见习记者 樊晓丽

“学生如果只发数据论文，是毕不了业的。”近日，北京师范大学教授、《数据快报(英文版)》(*Data Express*)副主编左西年接受《中国科学报》采访时说。数据的价值有目共睹，但在当下的科研评价体系中，这样“不被认可”的尴尬并非个别。

近日，中国科学院在数据期刊建设工作推进会上宣布，我国首本英文数据期刊《数据快报(英文版)》正式创刊，并将在年内陆续启动建设生态环境、海洋大气等领域的19种数据期刊，形成“1种综合旗舰刊+19种领域专业刊”的数据期刊集群。这一部署意味着科学数据开始从“附庸”走向“独立”。它既关乎科研评价体系的深层重构，也是对人工智能(AI)浪潮、数据主权竞争与科研范式革命三重变局的深切回应。

从“附庸”到“独立”

当前，AI大模型从拼算力、拼算法进入拼数据阶段。中国科学院计算机网络信息中心党委书记孙德刚指出，AI时代，科学数据已成为新型战略资源。高质量、可信、可溯源的科学数据是科技创新的“种子”。中国科学院组织创建的数据期刊集群，目标正是助力国家服务科技发展的新型基础设施能力建设。

这套基础设施的核心在于赋予数据独立的“学术人格”。通过规范的元数据、完整的方法学记录和独立标识号，科学数据得以摆脱“论文附件”的从属地位，成为可被独立检索、评价和引用的学术成果。这并非否定过去的成绩，恰恰是因为传统共享模式已触达能力边界。

中国科学院空天信息创新研究院研究员邱玉宝用一段亲身经历进行解释。2015年，他从地球观测组织(GEO)瑞士日内瓦秘书处回国，发现自己的硬盘里堆满了数据。“老有人追着我拷数据，每次我还要从头解释这数据怎么来的、怎么处理的、疲于应对。”他说，“我自己找起来都费劲，更别说明别人了。”

就在此时，邱玉宝得知中国科学院正在筹建科学数据银行，便主动联系，把最头疼的那批数据整理好、描述清、打上标签，成了这个平台的第一批用户。“数据藏在自己硬盘里，价值是死的；放到平台上可检索、可引用，价值才是活的。”

在中国科学院院士、《数据快报(英文版)》主编于贵瑞看来，数据期刊是鼓励科学数据开放共享、完善数据治理体系的重要抓手。数据期刊集群的创建恰恰填补了我国数据出版空白。

已有实践证明，数据一旦独立、共享，价值可以无限延展。20多年前，于贵瑞带领团队建立了中国通量观测研究网络，持续分享陆地生态系统观测数据。如今，这些数据已成为国内碳循环、水源涵养等领域的重要基石，支撑了数千项后续研究。

左西年对此同样感触颇深。10多年前，他在牵头共享脑影像样本时发表了3篇核心数据论文。其中一组30人每隔3天重复测量的脑部数据当时被视为“无用”，最终却被一个做物种研究的团队直接采用。今年4月，他们的研究成果发表于《科学》。

十年“造船”

一口气布局20本数据期刊绝非一日之功，背后是中国科学院长达10余年的系统性布局。

2015年，中国科学院创办我国第一本中文数据期刊《中国科学数据》。这本期刊像一块“试验田”，一步步摸索数据论文怎么写、怎么审、怎么评。《中国科学数据》编辑部主任孔丽华在推进会上表示，创刊11年来，这本期刊已组织40多个专题、发表超千篇数据论文、出版数据量近30太字节(TB)，积累了数

据出版全流程的政策与规范经验。

这揭示了数据出版的复杂性，涉及数据清洗、元数据抽取、存储格式转换、长期保存等一整套技术链条。如果没有前期经验，现在面对20种期刊集群的技术支撑将无从谈起。

不过，想要走向国际、掌握话语权，仅有一本中文期刊远远不够，还需要标准和平台的双轮驱动。

标准方面，中国科学院牵头制定了从数据引用、元数据规范到论文编写的多项国家标准，其中《数据论文出版元数据》和《数据论文编写规则》(《信息与文献编写规则第4部分：数据论文》)两项核心标准已提交至国际标准化组织。一旦通过，这将是 中国从国际数据标准“跟随者”向“制定者”迈出的关键一步，意味着未来的全球数据论文可能需要遵循“中国规则”撰写。

平台方面，中国科学院已建成投审稿、数据存储、资源标识、预印本等覆盖数据出版全链条的数字平台，且全部实现了国产自主可控。其中，数据存储平台已获国际主要出版商的广泛推荐，并被联合国教科文组织评为开放科学优秀实践案例。

科学出版社副总经理张凡说：“过去我们靠国际出版商的平台‘借船出海’，现在我们要造自己的船，驶向远海。”他同时强调，自主不等于封闭，是为了不被“卡脖子”；开放，是为了让中国数据成为全球公共产品。

实际上，数据期刊并非中国首创。爱思唯尔的《数据简介》和施普林格·自然的《科学数据》均创办于2014年。张凡坦言：“全球范围内，数据期刊仍属于相对新鲜的事物。这次集群化发布让中国有机会与全球并跑，甚至在某些领域实现超越。”

这份信心来自中国数据期刊集群背后的国家战略需求牵引和国家级数据基础设施托底。中国科学院计算机网络信息中心主任周园春介绍，11个国家科学数据中心、覆盖60多个一级学科的数据积淀，为数据出版提供了完全不同于商业化出版社的底层支撑。

最难一关

“数据论文发了，学校认不认？”中国科学院科技基础能力局副局长张韵向《中国科学报》记者坦言，这是当前较棘手的制度性障碍。

传统学科为何难以接纳数据论文？成都理工大学副教授杨雷解释说，日常数据采集通常带有明确的目的，“是为了验证我们的某一个观点或模型”。“发文章的主要目的是阐述学术观点、报道学术发现。”在这种科研传统下，数据被视为支撑学术观点的工具，而非独立成果。

即便如此，杨雷仍表示，“经过严格评审的数据论文发表还是非常有意义的，可能会给很多科学家带来启发和潜在的科学发现”。这种既认可数据价值又难以将其纳入学术产出体系的心态，在传统学科领域颇具代表性。

这样的认知变化带来的是科研评价体系的逐步转变。邱玉宝介绍，在他所在的院校，学生发数据论文是被认可的。“我们单位对数据成果的认可度正慢慢向传统期刊对齐。”

对常年驻守野外台站、大科学装置和观测网络中的数据采集人员而言，数据论文被认可，意味着即便不擅长撰写传统论文，他们的汗水也可以凝结为独立的、可被引用的学术产出。

“社会是有分工的。”于贵瑞说，“有人负责造砖，有人负责盖楼。你不能因为盖楼有显示度，就轻视造砖的人。”在他看来，数据科学家的价值应该被看见、被认可、被激励。

政策层面已出现破冰信号，中国科学院也率先迈步。据张韵透露，中国科学院正在制定指导性意见，推动将数据

论文纳入职称评定参考标准，强化数据论文作为学术成果的认定工作。

数据论文的勃兴也伴随着新的隐忧，防范数据造假仍是全球性难题。目前，国际上尚无任何期刊能对数据集的真伪做出结局验证，审稿人通常只能核查数据是否规范、流程是否合理、是否满足FAIR原则(可发现、可访问、可互操作、可重用)。数据的真实性更多依赖学术共同体的信任背书及审稿人的专业判断。

此外，新刊进入科学引文索引通常需要3至5年的观察期，且须在本领域排名前列。在此期间，如何让国内外学者认可并可主动引用这些数据成果，仍是《数据快报(英文版)》必须长期面对的生存挑战。

话语权之争

全球范围内，主要经济体已将高质量数据视为关乎国家安全的战略资源。中国作者发表的高水平国际期刊论文已占世界总量的39.2%，论文数量和被引用次数均位列全球第一，但其中绝大多数发表在 国外期刊上。论文在外、数据在外的“两头在外”局面至今未根本改观。

这种局面的代价是惨重的。据《2024年全球OA期刊与论文处理费监测报告》，2024年中国通讯作者在科学引文索引收录的OA(开放获取)期刊上发表的论文，文章处理费总支出约9.09亿美元，折合人民币约64.74亿元，八成以上流向了爱思唯尔、施普林格·自然等国际出版社。

“我们写了论文、贡献了智力成果，出版社还要收我们高额开放获取费，这在逻辑上是讲不通的，而且这个费用在科研经费中占了非常大的比例。”左西年表示，自主平台的建立不仅是出于安全考量，更是对巨额科研经费外流的纠偏。

张凡补充说：“尤其是在当前国际局势下，把科学数据留在自主平台上至关重要。”

这正是中国科学院下决心打造数据期刊集群的深层动因。如果不能掌握数据出版的话语权，中国在AI时代的科技竞争就将受制于人。

“数据共享不只是技术问题，更是规则主导权问题。”邱玉宝直言，在国际组织中推动数据开放时，最大的障碍往往不在于是否愿意共享，而在于“数据该怎么引用、怎么标识、怎么评价”。“如果标准是别人制定的，我们就要按照别人的规范走。现在推出自己的数据期刊集群，就是在搭建自己的学术舞台、建立能主导的话语体系。”

于贵瑞对此有更长远的期盼：“过去，中国科学家大量依赖国外的公开数据开展研究。未来，能不能让更多国际同行使用中国的数据做研究？我们要为世界提供可共享的公共数据产品，这也是数据强国的标志。”

“国际上没有哪个国家像中国这样把数据治理提到如此战略高度。”在邱玉宝看来，数据期刊集群在2026年的诞生“恰逢其时”，让数据具有“温度”、可以被感知。“我们不必永远做追随者。我国拥有庞大的科研群体，不仅是数据的生产者，也是数据的科学使用者。在数据治理和AI就绪这条新赛道上，中国完全有底气从并跑走向领跑。”

从一本中文刊到20本期刊集群，从依赖国外平台到全链条自主可控，中国科学院的数据期刊建设不是一次孤立的学术出版行为。它同时在做3件事：回应科研评价之困，让数据的价值被看见、被衡量、被承认；回应AI国家战略之需，为AI提供高质量、可信任的“燃料”；回应国际话语权之争，在数据成为战略资源的时代，把出版平台、评价标准和数据主权掌握在自己手中。

让AI学会“脑补”，给稀疏数据“补短板”

■本报记者 刁雯蕙



研究示意图。研究团队供图

大脑是高度复杂的精密系统，即使只看到物体的一部分，也能“脑补”其完整形状；即使场景信息不完整，也能通过连续的空间理解认识物理世界。那么，在人工智能(AI)快速发展的当下，智能系统能否像大脑一样，高效完成类似的重建工作？

近日，南方科技大学副教授王中锐团队联合复旦大学教授刘琦、刘明团队，中国科学院微电子研究所研究员尚大山团队等在《自然》发表研究，提出了基于忆阻器存算一体的高效神经场重建系统。该系统如同给AI装上“智能大脑”，在医学影像、三维视觉和动态场景的特定任务中，较传统图形处理器(GPU)能效提升数十倍。

该成果突破了传统计算框架中存算分离导致的功耗过大、数据缺失等瓶颈，有望为医学影像、增强现实(AR)/虚拟现实(VR)和具身智能的高效AI系统提供一条新的硬件实现路径。

那么，能否用更紧凑、更协同的计算方式呈现复杂世界，并用更高效的硬件完成计算？

由此，研究团队提出了一个核心思路——将“神经场”和“忆阻器存算”结合起来。忆阻器既可以存储信息，也可以参与计算。通过忆阻阵列，矩阵运算可以在存储单元内部完成，从而减少数据搬运，提高并行计算效率。根据这一特性，他们构建了基于40纳米、256Kb的忆阻器存算一体的神经场重建系统。

“这项研究历时4年，最困难的部分是让软件算法和物理硬件相互适配。起初我们用传统方法直接把模型部署到忆阻器阵列上，结果重建出来的图像非常模糊。”论文共同通讯作者刘琦回忆，团队花了很长时间思考，是硬件不行还是使用硬件的方式不对。

“不能把硬件当作‘完美的计算单元’来用，而应该让算法去适应硬件的不完美。”刘琦介绍，正是这种思路的转变，让他们提出了一种“硬件感知量化”方法，通过动态补偿误差的方法提高了计算的精度。

此外，团队还在软硬件等多个层面进行了协同创新。在算法层面，团队通过硬件感知的设计让算法变得更“轻”，从而适配边缘计算等计算资源有限的场景；在硬件层面，他们利用忆

针对这一挑战，研究团队提出了一种机制上区别于现有方法的非氧化还原骨架构编辑策略。该策略采用“开环—关环”的设计思想，在同一反应体系中依次完成路易斯酸促进的氧杂环丁烷开环胺化和分子内脱水环化两个关键步骤，实现了氧杂环丁烷向氮杂环丁烷的两步一锅法直接骨架构编辑。

与现有方法相比，新策略无需氧化还原活化，操作简便，具有广泛的底物适用范围和优异的官能团兼容性，并能够高效实现药物化学中具有重要价值的3,3-二取代氧杂环丁烷向相应氮杂环丁烷的直接骨架转换。此外，该策略还可直接应用于复杂生物活性分

阻器的“随机性”特点，生成高斯编码所需的随机矩阵，以帮助神经场获得更丰富的特征表达能力。通过这种软硬件多个层面的协同，团队最终实现了在器件低功耗下进行高精度的信号重建。

“正是在多学科深度交叉的合作模式下，团队逐步攻克了器件物理、集成电路设计、计算机体系结构、机器学习等方面的技术难题。”刘琦介绍。

较 GPU 能效提升数十倍

为验证该系统的性能，研究团队将其运用于医学影像、三维视觉和动态场景。

在三维CT稀疏重建任务中，该系统能够根据有限数量的CT切片恢复较完整的三维医学影像，相较GPU实现约23.5倍的折算能效提升，为低剂量、快速医学成像提供了潜在方案。在新视角合成任务中，该系统根据已有图像生成不同观察角度下的三维场景图像，相较GPU实现约21.0倍的折算能效提升，有望用于AR/VR、三维内容生成和机器人视觉等场景应用。研究团队还进一步评估了系统在动态场景新视角合成任务中的应用潜力，实现32.3倍的折算能效提升，说明该系统不仅适用于静态三维重建，也具有处理复杂时空信号的潜力。

审稿人评价道，该工作探索了一个非常有趣且有用的方向，在神经场重建的软硬件协同设计领域作出了有价值的贡献。

“这项研究的意义不仅在于提升了某一个任务的速度或能效，更在于提供了一种新的AI系统设计范式与计算范式。新型存算硬件不只是传统处理器的辅助，它或许能成为设计高效AI系统的重要组成部分。”论文共同通讯作者尚大山说道。

“利用该技术，未来的医疗仪器或许只需要采集少量CT切片，就能快速、低功耗地重建完整三维影像，从而缩短患者暴露在辐射中的时间。在AR/VR及具身智能领域，该技术有望使边缘设备在仅靠一块小电池供电的情况下，实时理解并渲染出复杂的三维空间。”论文共同通讯作者王中锐表示。

相关论文信息：<https://doi.org/10.1038/s41586-026-10646-w>

发现·进展

中国科学技术大学

提出概念独特的新策略实现四元杂环骨架精准编辑

本报讯(记者王敏)中国科学技术大学特任教授张海军团队发展了一种非氧化还原骨架构编辑新策略，实现了氧杂环丁烷向氮杂环丁烷的直接骨架构转换，为优势杂环之间的快速互酸促进的氧杂环丁烷开环胺化和分子内脱水环化两个关键步骤，实现了氧杂环丁烷向氮杂环丁烷的两步一锅法直接骨架构编辑。

与现有方法相比，新策略无需氧化还原活化，操作简便，具有广泛的底物适用范围和优异的官能团兼容性，并能够高效实现药物化学中具有重要价值的3,3-二取代氧杂环丁烷向相应氮杂环丁烷的直接骨架转换。此外，该策略还可直接应用于复杂生物活性分

化等多种疾病损伤；而C5aR2的分子作用、信号特征长期存在争议，是补体免疫领域亟待破解的关键科学问题。该研究利用高分辨率冷冻电镜解析了C5aR2复合物三维结构，明确其区别于经典G蛋白偶联受体、依赖β-arrestin通路的独特信号模式，对比阐明C5aR2与C5aR1在构象、配体结合、信号偶联上的核心差异，从分子层面解释了二者信号偏向性的内在机制。

在此基础上，团队通过结构指导的理性设计得到高选择性C5aR2完全激动剂ZQ105。借助该特异性工具分子，研究证实激活C5aR2可调控中性

中南大学湘雅二医院等

推翻“沉默受体”认知开发高选择性靶向激动剂

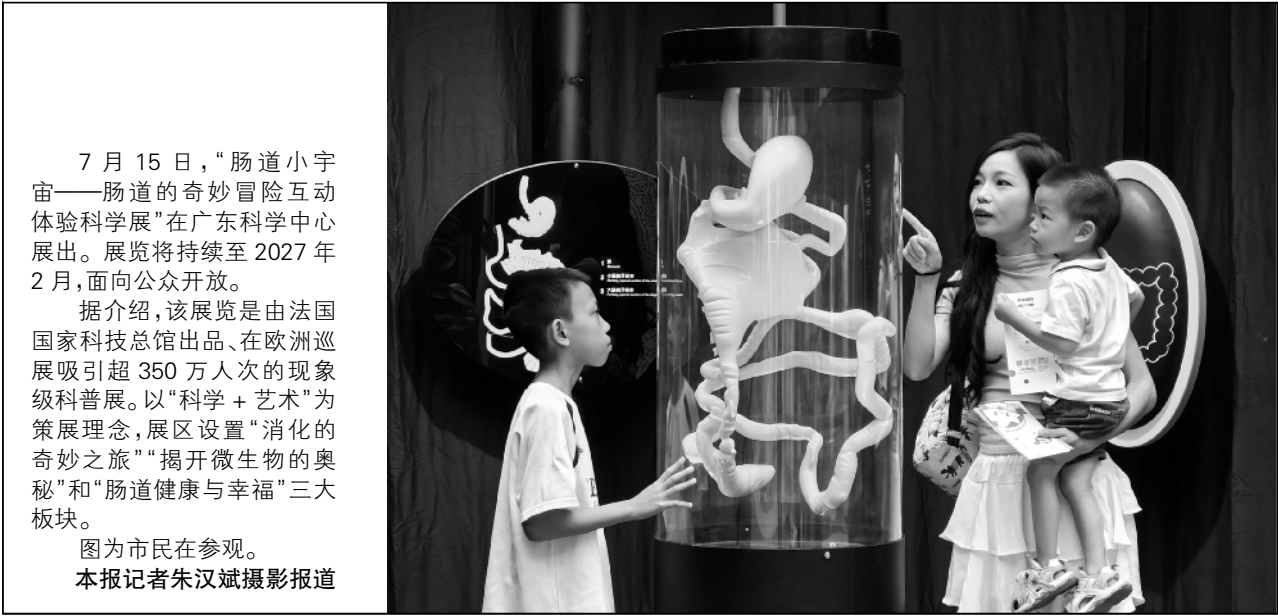
本报讯(记者王昊昊)中南大学湘雅二医院教授吴海竟团队联合中国医学科学院皮肤病医院教授陆前进团队、浙江大学教授张岩团队及四川大学教授邵振华团队，揭示了补体受体C5aR2配体识别、非典型信号转导的分子机制，并开发出高选择性靶向激动剂，为炎症相关疾病诊疗提供全新理论基础与工具分子。相关成果近日发表于《细胞研究》。

补体系统是人体固有免疫核心组分，C5a作为强效促炎介质，可通过C5aR1、C5aR2两类受体调控炎症。现有研究多聚焦C5aR1通路，证实其与脓毒症、红斑狼疮、血管炎、肾纤维

粒细胞炎症活化，改变细胞表面功能分子表达，促进炎症因子分泌，并参与细胞内容与溶酶体系统重塑，推翻了以往将C5aR2视作“沉默受体”的认知，证实其可独立放大炎症反应、重塑免疫细胞功能。

据介绍，系统性红斑狼疮、自身免疫性皮肤病、脓毒症、肾纤维化等疾病均存在补体过度激活、中性粒细胞异常浸润损伤。该成果厘清了C5aR2在炎症进程中的作用机制，同时提供特异性靶向工具，为研发C5aR2靶向药物、双受体联合干预方案开辟了新路径。

相关论文信息：<https://doi.org/10.1038/s41422-026-01273-1>



7月15日，“肠道小宇宙——肠道的奇妙冒险互动体验科学展”在广东科学中心展出。展览将持续至2027年2月，面向公众开放。

据介绍，该展览是由法国国家科技总馆出品、在欧洲巡展吸引超350万人次的现象级科普展。以“科学+艺术”为策展理念，展区设置“消化的奇妙之旅”“揭开微生物的奥秘”和“肠道健康与幸福”三大板块。

图为市民在参观。本报记者朱汉斌摄影报道