

# 建好科技语料“粮仓”，给 AI 喂“好粮”

■本报记者 赵广立

人工智能赋能科学研究(以下简称 AI4S)作为人工智能“皇冠上的明珠”,正在引领新一轮科研范式变革,加速各领域科技创新突破。然而,作为科研新范式必备工具,其发展高度依赖科技语料的有效供给。

近日,在 2026 世界人工智能大会期间,由中国科学院文献情报中心牵头,联合国内科学数据、科技文献领域近百家优势单位和覆盖全学科领域专家共同建设的中国科学院“散仓·科技语料库”(以下简称散仓)正式发布,标志着我国支撑人工智能高质量发展的新型基础设施建设和取得突破性成果。

“高质量科技语料库已经成为人工智能领域战略性稀缺资源。”中国科学院文献情报中心主任曲建升表示,建设散仓的目的,就是要建好科技语料“粮仓”,给 AI4S 喂“好粮”。

## 建成国家级语料库的“底气”

什么是科技语料?据介绍,科技语料是对科学研究活动中产生的多模态科学数据与科技文献等进行清洗整理、解构重组、关联融合,提升语义密度与知识含量后形成的高价值智能就绪(AI-Ready)数据。

曲建升说,建设国家级的科技语料库对于突破 AI4S 语料供给瓶颈、夯实人工智能数据底座,保障科技创新自主可控、抢占全球 AI4S 发展制高点具有重要意义。然而,由于数据模态多、知识密度大、质量要求高、加工治理难等

特点,科技语料建设具有较高的门槛与壁垒。

目前,中国科学院拥有国内最大的科学数据产出与储备体系、国内最全的科技文献资源建设与保障体系,具备由高水平、全学科的科研队伍形成的专业化语料加工治理能力,以及广泛的科学应用场景。

正是依托中国科学院的建制化、体系化优势,散仓才有了建成国家级语料库的“底气”。以数据供给为例,其来源于超过 320PB 的科学数据、1.5 亿篇科技论文、1.2 亿项发明专利和 500 万本科技图书,覆盖文本、图像、图表、公式、影像、化学式、波、谱、场等多种数据形态。

“中国科学院牵头建设散仓,体现了国家战略科技力量主力军服务国家重大战略需求、引领全球科研范式变革的使命担当。”曲建升表示,散仓的建成,对筑牢国家人工智能发展数据底座、实现高水平科技自立自强,具有不可替代的基础性、战略性价值。

## 不是数据的简单堆砌

散仓之名,源自古代位于河南荥阳的粮仓,是秦统一六国后建成的全国最大粮仓,承担着汇聚天下粮储、保障国家供给的战略使命。

如今,散仓致力于打造国家级科技语料“粮仓”。

曲建升告诉《中国科学报》,科技语料库不

是数据的简单堆砌,而是一项体系性极强、专业度极高、多单位协作的系统性工程。

目前,散仓构建了“1+7+N”总分库模式,即 1 个全学科数据知识融合的国家级总库、7 个专业化深度提炼的领域分库与 N 个场景化关联的应用子库,并通过制定科技语料描述、加工、评价、服务及运维的全流程标准,形成了数学物理、化学化工、天文空间、地球科学、生命科学等全学科覆盖的“内容就绪、计算就绪、可信就绪”的科技语料供给体系。

“散仓中高知识含量、高语义密度的语料,可以补齐大模型领域知识短板、校准逻辑推理偏差,使其输出符合科学基本规律与数理逻辑,具备领域专业知识与长程思考能力,提升多模态理解和跨领域交叉能力。”曲建升说,为给用户提供更科学严谨、更智慧专业、更灵活适配、更规范可信的语料服务,建设团队创造性地提出基础语料、特征语料、知识语料三阶知识精炼模式,大幅提升语义密度和知识密度。

曲建升说,正是依托这套全栈自研的语料库技术体系,散仓实现了从原始数据到智能就绪语料的高效加工转化,为下游应用奠定了坚实基础。

## 场景驱动,边建边用

语料库的价值体现在共建共享、有机生长。作为面向 AI4S 新型基础设施的国家级科技语料库,散仓设计之初便立足于语料共建共

享、可持续化运营,赋能科学研究和产业发展。

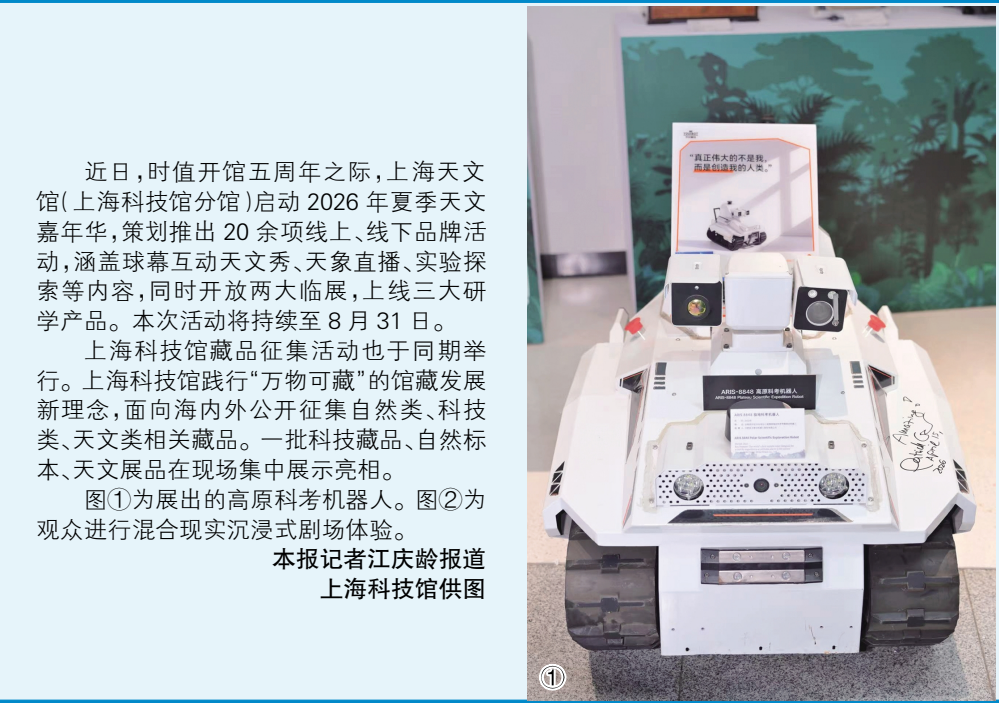
据介绍,秉承“场景驱动、边建边用”的理念,散仓以赋能科学研究、人工智能应用与产业创新为核心使命,在许多场景的服务中已经显现成效。

面向科学发现场景,散仓相关语料已多批次供给“磐石”科学基础大模型,助力大模型科学推理能力显著提升;同时,对智能化工大模型等多领域学科模型按需供给,有效支撑数学物理、化学化工、天文空间、地球科学、生命科学、前沿科学等领域取得新进展。

面向智能应用场景,散仓已经开始为蚂蚁医疗大模型提供支撑服务,助力蚂蚁阿福为医生、患者提供更为严谨规范的科学输出;同时,散仓团队正在启动对豆包、千问等通用大模型及科大讯飞等领域应用大模型的语料供给工作。

面向产业创新场景,散仓正针对商业航天、低空经济、智能制造、生物医药等领域开展语料加工服务,下一步将支撑元器件选型、超微病理诊断、高速列车外形设计等百余个行业应用。

谈及未来发展,曲建升表示,散仓将坚持科技语料“共建、共享、共用”原则,强化高质量科技语料加工能力,丰富 AI4S 创新应用场景,探索公益与商业结合的语料运营模式,培育科技语料繁荣创新生态,赋能新质生产力跨越发展,为我国乃至全球科技创新与智能经济发展贡献科技语料库的支撑力量。



近日,时值开馆五周年之际,上海天文馆(上海科技馆分馆)启动 2026 年夏季天文嘉年华,策划推出 20 余项线上、线下品牌活动,涵盖球幕互动天文秀、天象直播、实验探索等内容,同时开放两大临展,上线三大研学产品。本次活动将持续至 8 月 31 日。

上海科技馆藏品征集活动也于同期举行。上海科技馆践行“万物可藏”的馆藏发展新理念,面向海内外公开征集自然类、科技类、天文类相关藏品。一批科技藏品、自然标本、天文展品在现场集中展示亮相。

图①为展出的高原科考机器人。图②为观众进行混合现实沉浸式剧场体验。

本报记者江庆龄报道  
上海科技馆供图



# 是党员,也是空间应用系统的“托底人”

■刘如楠

得知返航推迟消息的那一刻,大家的心一下子悬了起来。

这是酒泉卫星发射中心,神舟二十号载人飞船计划于 2025 年 11 月 5 日返回,4 只“小鼠航天员”也一同回到地面。然而,就在前一晚,团队接到了返回任务推迟的通知,推迟时间不定。

“3 天、5 天、9 天……”最终,小鼠的“出差”时长由原计划的 5 天骤增至 14 天。

最大的风险是断粮。一旦断粮,小鼠性命堪忧。“得知消息后,我们团队立刻开始制定多种补给方案,推演可能结果,最终保障了小鼠的生命健康。”中国科学院空间应用工程与技术中心(以下简称空间应用中心)系统总体研究室党支部书记杜昌帅说。

在这次空间实验任务中担当总体协调角色、发挥关键作用的,是空间应用中心系统总体研究室。近日,该研究室党支部获得中国科学院先进基层党组织荣誉称号。

## 鼠粮补给:让小鼠活着回来!

时间拉回到 2025 年 11 月 4 日晚,大家刚刚吃过晚饭,突然意外消息传来。

神舟二十号飞船舷窗玻璃出现裂纹,疑似遭受空间碎片撞击。安全起见,航天员及“小鼠航天员”推迟返航。

“第一次上天,小鼠本就紧张,推迟返回,能不能撑住?”“粮食最多只能吃 10 天,没有余粮怎么办?”“能调粮吗,从哪里调”……问题一个接一个抛出来,没有人敢贸然回答。

紧急会议连夜召开,发射中心、指挥中心、实验室的灯光亮了一晚。“让小鼠活着回来”成了大家共同的目标。

返航时间未定,系统总体研究室的成员需要联合实验动物专家,评估所有返回时间窗口下的小鼠安全健康问题。

“我们迅速联系了航天员系统,看能不能让小鼠吃航天员的食物。玉米、豆浆、饼干……大家考虑了很多种食物,一旦列为备选,地面验证试验立刻启动。只有在地面验证过后,才敢让空间站的小鼠尝试。”空间应用中心系统总体研究室工程师金雪娜说。



工程师在工作,左一为杜昌帅。

空间应用中心供图

反复测试后,团队最终确定了合适的豆浆配比,既能满足小鼠的营养需求,又不至于因太过黏稠而堵塞饲养装置的水嘴。

有人可能会觉得,鼠粮少了、水少了,加一些不就好了?事实是,“天上无小事”。在地面上进行补给轻而易举,而在空间站的特殊饲养装置与微重力环境下,任何微小的操作都会造成一系列后续影响,特别是安全问题。

而系统总体研究室,正是系统规划各项科学实验、评估把控方方面面细节,最终保证实验顺利开展的“把关人”。

## 月面探测:既要一起走,又要走得快

如果说鼠粮补给任务是系统总体研究室在空间实验实施阶段面临的重大考验,那么在载人月球探测任务中,从前期规划论证阶段开始,团队便最大化地发挥起了统领协调作用。

当 2030 年载人登月任务确定后,一系列前期试验任务就随之开始。其中一项,便是两台探测器无人飞行、月面自主探测。

其难点在于,国内首次进行两台探测器并行探测,探测器要短时间内在月面长距离行走,完成相关地质探测任务。这意味着,两台探测器“既要一起走,又要走得快”。

“以前一台探测器在月面行走,地面只需

测控一个目标。现在两台同时走,理想情况是由地面测控两个目标。但由于分给空间实验的地月通信带宽有限,无法实现同时对两个目标的测控,这是我们面临的重大难题。”空间应用中心系统总体研究室工程师李雨说。

无法实现同时行进,团队自然而然地想到“先富带后富”,或许可以实现一个先走、一个后走。

很快,这个方案就被否决了。“比如,前一台探测器走了 100 米后停下来,等着第二台走过来。在这个时间段,地面就失去了与第一台的联系,它是否发生意外也不得而知。”李雨说。

在汽车自动驾驶领域,跟随前车的功能必不可少。这是指通过激光雷达、车载摄像头等持续探测前方车辆、距离、相对速度等。

受此启发,团队提出了两台探测器“一个自主行进、一个自主跟随”的方案。如此一来,后面的探测器不需要实时分析周边环境信息,只需要判断与前方探测器的相对距离、速度等信息,便能实现稳定行走。

“系统方案确定后,系统总体人还要将方案自上而下分解,并抓总负责优势科研单位、工程单位的择优工作。我们争取把国内某方面最有优势的团队凝聚在一起,肩负起国家交给我们的重大科技任务,顺利完成重大使命。”杜昌帅说。

给大家托底:尽一切努力把科学载荷送上天

遇到一些因种种原因不能如期推进的项目,系统总体研究室的成员常常急得抓耳挠腮。

一次,某任务立项,需要多家单位联合攻关。在系统总体研究室的抓总下,多次技术协调会如期推进,敲定了相关总体技术文件。然而,就在大家按图索骥时,有一家单位却反馈

了一些技术问题,让原本紧张的研制时间变得更加局促。

“没办法,只能挨个找其他承研单位协调机、电、热、信息、软件等接口条件。牵一发而动全身,要在种种限制条件下获得最优解,只能靠我们逐个沟通技术细节。”李雨说。

时间紧、任务重,着急的时候,有的人真想甩手不管。“某个项目要求特别高,系统总体研究室也管得比较细,有的单位就说‘怎么螺丝刀的型号也要管’。这让人特别委屈。”金雪娜说。

每当这时,团队里的“老将”、系统总体研究室原党支部书记、研究员赵黎平的话就会回响在大家耳边:“系统总体工作做好了,就像打开水龙头会有水,打开开关灯会亮。润物细无声处,是稳稳的托举。”

大家互相排解、互相鼓励过后,消极的情绪也没了大半。金雪娜说:“我们空间应用系统作为载人航天工程的系统之一,国家太空实验室高质量高效运行、重大科技成果产出是我们的核心目标,我们必须为国家太空实验室项目争取好的资源条件,为项目的成功实施尽一切努力。”

乍一看,项目中的每一项具体工作都不由系统总体人完成,但他们的身影无处不在:立项前期要论证、评审、遴选;立项之后要制定系统方案,明确产品状态以及多方协调测试,“给各个零部件都抹上润滑油”;载前上了天,要随时盯着;样品返回地面,要稳稳交付……

这样的工作做久了,团队里的年轻姑娘甚至有了“当妈”的感觉。“科学载荷就像我们的孩子,我们扶着走、托着走、抱着走,尽一切努力把它们送上天。”金雪娜说。

整个研究室人员并肩作战,总有让人暖心的时刻。比如,故障发生时,不论需要谁的支持,一个电话就能把他召唤回来;配合载人飞船发射任务时,在发射台架下成立临时党支部,对着党旗齐声宣誓;各个载荷调试完成时,测试大厅里响起经久不息的掌声……

正是许多这样的时刻,一次又一次给了团队成员坚持下去的力量。

“系统总体研究室就是给大家托底的,不论发生什么,我们都在,要确保任务万无一失。”杜昌帅说。

# 热射病善于“伪装”老人更需小心

本报讯(记者朱汉斌 通讯员房诗婷)每年大暑前后,医院急诊科都会迎来一批特殊的“中暑”患者——他们并非户外工人或烈日下的运动员,而是未出门的居家老人。对此,家属往往很困惑:明明没出门,怎么就中暑了?

7 月 21 日,中山大学孙逸仙纪念医院急诊科主任医师方向韶在接受采访时表示,公众口中的“中暑”,在医学上涵盖一组由轻到重的症候群。热痉挛表现为大量出汗后电解质丢失引发肌肉痉挛,多见于运动或体力劳动后;热衰竭系因大量失水、失盐导致循环血量不足,出现头晕、乏力、大汗、恶心甚至晕厥状况,及时补水 and 休息多能自行恢复;热射病最为严重,此时体温调节机制彻底失灵,核心体温急剧升高,常并发多器官损伤,危及生命。

值得注意的是,热射病存在两种类型。劳力型热射病多发生于青壮年,诱因明确——高温下的剧烈体力活动,起病急骤,皮肤湿热、大量出汗。随着高温防护意识增强,城市中此类病例已相对少见。而经典型热射病恰恰相反,好发于老年人和体弱慢性病患者,无需任何剧烈运动,仅被动处于高温、高湿环境中,他们的产热与散热便悄然失衡。经典型热射病起病隐匿,可在一至两日内缓慢进展,皮肤干热、少汗甚至无汗,病死率达 30%至 70%,远超劳力型热射病。在急诊科,这类患者占中暑就诊的绝大多数。

“经典型热射病是一种‘被动’发生的疾病。老人没有主动走向高温环境,只是在自己家里,不知不觉就被热浪击倒。”方向韶说,经典型热射病的危险之处在于其善于“伪装”,常被误认为其他疾病而错失救治时机。临床中,老年热射病常以脑卒中面貌出现——言语含糊、一侧肢体无力、步态不稳;或表现为心脏病症状——胸闷、心悸、血压下降甚至休克;或类似感染或脓毒症——发热、精神萎靡、化验指标异常;有时仅表现为“人有点怪”——答非所问、烦躁、嗜睡,而这恰恰是早期意识障碍的信号。

《中国热射病诊断与治疗指南(2025 版)》有一项重要更新——核心体温超过 40 摄氏度已不再是诊断热射病的绝对前提。部分患者体温尚未达到此水平,却已发生多器官损伤,不能因“体温没那么高”就轻易排除热射病。

热射病为何对老年人特别危险?方向韶解释,人体散热依赖辐射、对流、传导和蒸发 4 种途径。而老年人的汗腺退化、血管调节能力下降、口渴感迟钝、心肺储备有限,加之如果常年服用  $\beta$  受体阻滞剂、利尿剂等药物,散热功能几乎全面受损。一些城市夏季湿度大、风力小,即便室外温度不是很高,一间没有空调、通风不良的老旧房间的室内温度也可能在数小时内升至危险水平。

方向韶强调,热射病救治的黄金原则是降温第一,转运第二。施救者应立即将老人转移至阴凉通风处或有空调的室内;脱去衣物,用冷水擦拭全身并配备风扇降温,冷水浸泡效率最高,同时重点冷却颈部、腋窝、腹股沟等大血管部位;立即拨打 120,等待过程中持续降温。绝不可“等等再看”,热射病进展迅速,每延误一分钟,器官损伤便加重一分。

## 发现·进展

山西医科大学 - 华大未来医学联合研究中心

# 揭示过敏性鼻炎核心致病机制

本报讯(记者刁雯蕙)近日,山西医科大学 - 华大未来医学联合研究中心联合基因组多维解析技术全国重点实验室,成功构建鼻黏膜组织单细胞组学图谱,揭示了上皮与基质细胞互作失调是过敏性鼻炎的核心致病机制。研究人员还开发了可预测过敏性鼻炎患病风险与临床指标的人工智能模型 scMARIA,为个性化治疗提供了全新思路。相关成果发表于《过敏与临床免疫学杂志》。

鼻黏膜组织是人体免疫系统的第一道关键防线,在抵御病原体入侵与维持免疫耐受之间维系精密的动态平衡。然而,在过敏性鼻炎患者中,这种平衡被打破了。该疾病涉及鼻黏膜中多种细胞类型及交互网络的全景式失调,传统研究手段难以系统解析底层机制。

科研人员对 25 例过敏性鼻炎、14 例非过敏性鼻炎患者的鼻黏膜组织进行了单细胞转录组测序,并对其中 24 例患者的样本补充了单细胞表观组测序,累计获得约 102.4 万个高质量单细胞组学数据。分析发现,两组患者的基底细胞、成纤维细胞亚型在基因表达和染色体开放调控层面存在显著差异,提示鼻炎发生后两类细胞在转录与表观调控层面出现了明显异常。

研究发现,过敏性鼻炎的致病机制不仅涉及免疫异常,还源于上皮细胞与基质细胞的失调。Club 细胞是鼻腔上皮屏障的一种重要细胞。在健康状态下,Club 细胞数量稳定、分化顺畅,负责不断补充受损的上皮屏障。过敏性鼻炎患者鼻腔中 Club 细胞数量减少且分化受阻,这些未成熟细胞转而释放促炎及促组织重塑因子,异常程度与症状严重程度相关。

基质细胞负责支撑鼻腔组织结构,调控局部炎症反应。研究发现,活化巨细胞和 CXCL2 成纤维细胞两类基质细胞亚群的比例虽然下降了,却高表达 POSTN、FADS1 等促炎重塑基因,可能加剧上皮损伤与组织重构。两者共同构成促炎微环境,推动疾病进展。

此外,团队还发现,过敏性鼻炎患者鼻腔内的上皮细胞与基质细胞之间的细胞对话异常活跃,涉及与炎症和组织重塑相关的通路,形成了“损伤 - 炎症 - 更严重损伤”的恶性循环。

相关论文信息:  
<https://doi.org/10.1016/j.jaci.2026.06.020>