

我国智能算力规模居世界前列

基于最新人工智能理论、采用领先计算架构的智能算力，正在为人工智能发展提供有力支撑。

党的二十届四中全会通过的《中共中央关于制定国民经济和社会发展第十五个五年规划的建议》提出，“加快人工智能等数智技术创新，突破基础理论和核心技术，强化算力、算法、数据等高效供给。”

工业和信息化部数据显示，截至今年6月底，我国在用算力中心机架总规模达1085万标准机架，智能算力规模达788百亿亿次/秒，存力规模超过1680艾字节，已发布1509个大模型，在全球位居前列。同时，智能算力应用走向深入，与各行业耦合作用不断增强，成为数字经济新底座。当前，智能算力应用主要体现在哪些方面？算力新产品、新服务模式怎样为高质量发展提供有力支撑？记者进行了采访。

驱动科研范式变革，加速原始创新成果产生

在西湖大学非编码核酸生物学实验室，研究员申恩志正在开展非编码RNA（核糖核酸）的作用机制和生物学功能研究。“非编码RNA约占人类转录组的98%，它不仅参与生命体的各种基本生命过程，而且与很多重大疾病的发生密切相关。”申恩志介绍。

非编码RNA数量庞大，而且在生命体里是高度动态的，可以跟很多其他生命分子相互作用、相互调节，如果单纯采用传统实验方式研究，需要耗费大量时间和精力，也很难分析其中复杂的调控关系并找出规律。如今，申恩志团队利用智算技术高效解析测序数据，解决了以前单靠实验解决不了的问题，大大加速了研究进程。

近年来，西湖大学在浪潮信息等企业助力下打造了领先的算力平台，支撑科学家开展跨学科研究。申恩志认为，人工智能与生命科学尤其是计算生物学的跨学科研究，显著提高了科学研究的速度和准确性，让很多新的发现从不可能成为可能。

依托智能算力实现高效模型训练与推理，极大加快了人工智能在科研领域的应用创新。国内高校院所和算力企业也在加快产学研合作步伐，推动智能算力在科研领域落地。

南开大学智能计算系统研究室在眼底图像血管分割研究中，借助高效人工智能计算开源框架，将单张眼底图像的推理速度提升了2.4倍，从而推动眼底图像分析技术的临床应用；上海人工智能实验室和临港实验室、上海交通大学、复旦大学等紧密协作，仅用2个月就发现、验证了2个癌症新靶点；鹏城



图片来源：网络

实验室通过中国算力网整合20余座城市的算力节点，依托“鹏城云脑Ⅱ”服务数万名科研人员……

在清华大学人工智能研究院常务副院长孙茂松看来，人工智能有望成为驱动科学范式革新的重要工具。在人工智能驱动的科研范式变革中，要善于利用智能算力发现“根节点”问题，如大模型预测蛋白质结构等，“解决它，就可能给该领域带来根本性变化”。

成果落地加快，适配多样化应用场景

国务院印发的《关于深入实施“人工智能+”行动的意见》提出，“强化智能算力统筹”“推动智能算力供给普惠易用、经济高效、绿色安全”。

位于内蒙古呼和浩特的伊利现代智慧健康谷，牧场里的奶牛拥有专属的“健康档案”。基于阿里云提供的人工智能算力以及飞天操作系统，伊利构建了“一云多芯、一云多算”的智能算力基础设施，为人工智能应用提供了稳定、弹性的部署环境，并且将决策算法和生成式大模型融入生产和供应链领域。

“靠人工观察奶牛健康状态，效率低且主观性强，现在通过计算机视觉识别奶牛眼部特征，我们能够对奶牛健康进行实时监测与精准饲喂管理，有效提升了产奶质量。”伊利集团数字科技中心数字技术总监程国强介绍，在伊利智能工厂，智能算力赋能自动化生产线，不断提升生产和管理效率。

“目前，伊利已开发超800个智能体。大模型覆盖订单履约、库存周转、物流时效等70%的供应链场景，显著降低了原辅料临期、积压及缺货风险。”阿里云解决方案高级总监贾朝辉表示，未来他们还将打造“智算云”，并结合牧草种植、奶牛健康等专业数据和行业知识，共同开发农牧行业大模型。据

介绍，自2023年开源第一款模型以来，阿里通义大模型在全球下载量突破6亿次，衍生模型突破17万个。

随着高速动车组速度等级不断提升，气动外形设计成为研发的关键。中车集团以既有仿真、试验数据为基础，构建高速动车组气动载荷标准数据库，依托百度飞桨深度学习平台以及飞桨科学计算套件，构建了空气动力学智能化仿真大模型。

中国中车集团科技质量与信息化部数字化创新处副处长、人工智能办公室推进组组长陈鉴介绍，与传统仿真方法相比，大模型实现仿真周期由超算资源的天级缩短至单机显卡的10秒级，整体仿真效率提升30倍以上，结果误差小于5%，推动我国高速列车外形设计从经验驱动向数据驱动转型，推动高速动车组外形的智能优选和快速优化。

百度基于文心大模型与飞桨深度学习平台打造的人工智能基础技术底座，正在助力千行百业提效降本。百度首席技术官王海峰介绍，飞桨支撑超大规模自适应分布式深度学习训练，万卡集群上大模型有效训练时间占比达到98%，通过大模型压缩、推理、服务部署全链条协同优化，实现低时延、高吞吐的大模型推理。此外，飞桨持续积极开展硬件统一适配，为算力多元化提供了有力保障。

产学研协同，新产品新模式不断涌现

近日，浪潮信息发布了面向智能体优化的人工智能计算系统。其中，基于元脑SD200超节点人工智能服务器，DeepSeek R1大模型词元（模型输入输出的基本单位）生成速度仅需8.9毫秒；最新发布的元脑HC1000超扩展人工智能服务器，使词元生成成本大幅降低。

超节点服务器是在有限物理空间内，通过极低时延的总线级互联技术和极简互连架构，打造一个逻辑统一的紧凑计算单元，整个系统可以像一台超大号计算机那样运转。

“开发超节点的关键是计算架构的创新，即跟随应用和算法发展方向，以应用为导向，以系统为核心，有效破解智能算力发展的瓶颈问题。”在浪潮信息首席人工智能战略官刘军看来，人工智能的大范围应用不断推动算力高速增长，发展智能算力要克服系统规模逼近工程极限、能耗过高等挑战，有效平衡投入与产出。

在业内专家看来，通过软硬件协同设计与深度优化，持续推动人工智能计算架构的创新与突破，才能进一步促进大模型、智能体等人工智能技术与实体经济的深度融合，让人工智能成为赋能千行百业的创新力。

2024年，在北京市科委中关村管委会组织下，浪潮信息等多家芯片、系统、大模型开发、应用创新等领域的企业和机构成立了“超节点算力集群创新联合体”，相关单位紧密协作，在超节点互连协议、系统研制、标准制定、应用部署等方面取得显著进展。

作为参与方之一，太初（无锡）电子科技有限公司积极与科研机构、运营商、算力服务商等合作打造智能算力解决方案，其推出的高密液冷智算集群已服务200余家高校企业，支持超1200千万亿次/秒浮点运算算力需求。“智能算力的创新突破，离不开产学研协同和人才支撑。”公司首席产品官洪源表示，他们与湖南大学、南京邮电大学、苏州大学等成立联合研发团队共同进行科研攻关，此外还构建了基于自研人工智能加速卡、软硬件协同的教学实训平台，培养了一大批年轻、有活力且具备实践能力的人才。来源：人民日报