

北京市首批人工智能行业大模型应用案例发布

场景驱动大模型迭代高速演进

■ 本报记者 张伟

6月27日,北京市科委、中关村管委会举办北京市人工智能策源地引领推介活动的首期活动——北京市首批人工智能行业大模型应用案例发布,发布北京市首批10个行业大模型典型应用案例。

据悉,北京市科委、中关村管委会正在持续征集北京市行业大模型典型应用案例和场景需求,编制北京市大模型技术产业创新图谱并将适时发布,加强对大模型行业应用的宣传推广。鼓励更多垂直行业用户积极开放场景,提供行业知识数据,联合人工智能创新企业共同开展行业大模型技术及应用研发,加速推动技术迭代与产业创新发展。同时从算力、数据、人才、投资等多个方面落地一批支持措施,持续做好大模型相关技术研发和产业培育工作,不断集聚创新资源,优化产业生态,加快建设具有全球影响力的人工智能创新策源地。

率先实现落地应用

在过去的半年中,以GPT-4为代表的大模型技术引领了新一轮全球人工智能创新热潮。国内外大模型科研高速迭代,各类模型层出不穷,模型性能不断提升,创新创业蓬勃兴起,展现了惊人的演进速度。同时,大模型赋能行业应用正在蓬勃兴起,基于大模型的泛化能力和涌现特性,其在越来越多垂直行业场景上初步展现了惊艳的性能表现和巨大的发展潜力,得到了各行业高度关注,场景驱动成为大模型技术创新的重要模式,有望形成新的产业业态和经济增长点。

北京市科委、中关村管委会二级巡视员刘航介绍,北京是国内大模型创新基础最好、人才团队最集中、研发能力最强、产品迭代最活跃的地区。一大批一线科学家和创业者集聚北京,涌现出文心一言、ChatGLM等一批具有代表性的大模型产品。科技部新一代人工智能发展研究中心发布的

《中国人工智能大模型地图研究报告》显示,我国10亿参数规模以上的大模型已发布79个,其中北京发布的数量最多共38个,占全国近一半,大模型学者指数、模型开源数量和影响力等指标均位居全国第一。

为进一步落实北京市《加快建设具有全球影响力的人工智能创新策源地实施方案》和《通用人工智能创新发展若干措施》,加快人工智能场景建设相关举措落地实施,北京市科委、中关村管委会瞄准国家重大战略、社会民生关切和垂直行业场景需求,把握当前阶段大模型技术演进趋势和行业应用特点,聚焦城市治理、医疗健康、科学研究、智慧金融、智慧生活、智慧城市等重点领域,组织专家凝练并发布现阶段技术创新性相对较强、成熟度相对较高、应用前景广阔的首批10个行业大模型典型应用案例。

典型场景多点开花

据介绍,此次发布的10个案例均由技术供给方与场景需求方联合研发,充分展现大模型技术能力,深度结合不同行业需求特点,在帮助行业用户“降本提质增效”的同时,开发新功能、形成新产品、拓展新应用,推动人工智能技术与行业更加深度融合。

据北京百度网讯科技有限公司百度智能云AI产业资深解决方案架构师许磊介绍,百度集团和国网智能电网研究院有限公司共同开发的基于电力行业NLP大模型的设备运检知识助手示范应用,在智慧能源领域,基于百度“文心”系列大模型及领先的深度学习、知识图谱增强等技术,构建了千万级电力文本样本库和电力行业知识图谱,联合训练电力行业NLP大模型,在电力专业分词、电力营销敏感实体识别F1指标方面,分别提升9.27%、13.28%,达到92.376%和94.947%,显著增强了设备及电网运营的自动化、智能化水平。

北京智谱华章科技有限公司CEO张鹏介绍,由北京智谱华章科技有限公司和北京中医药大学东方医院共同开发的数字中医大模型示范应用,在智慧医疗领域,基于智谱华章高精度千亿中英双语稠密模型“GLM-130B”,面向中医领域名医经验挖掘整理需求,构建数字中医服务平台,探索高危肺结节人工智能临床诊疗和临床评价研究等解决方案,实现中医临床经验的智慧化复制新模式。

云知声智能科技股份有限公司创新事业部副总裁郭凡介绍,由云知声智能科技股份有限公司和北京友谊医院共同开发的基于山海大模型的门诊病历生成系统示范应用,在智慧医疗领域,基于云知声700亿参数规模的自研“山海”大模型,结合前端声音信号处理、声纹识别、语音识别、语音合成等全栈式智能语音交互技术,预计可提升医生的电子病历录入效率超过400%,节约单个患者问诊时间超过40%,提升医生门诊效率超过66%。

其他技术供给方也介绍了应用案例情况,其中包括由中国科学院自动化研究所和中铁建设集团有限公司共同开发的面向建筑领域多模态行业大模型示范应用,由中关村科学城城市大脑股份有限公司和科大讯飞(北京)有限公司共同开发的自主可控城市大脑大模型示范应用,由北京科学智能研究院和中国航发北京航空材料研究院共同开发的覆盖元素周期表原子间势函数预测训练模型示范应用,由第四范式(北京)技术有限公司和北京中关村银行股份有限公司共同开发的多模态智慧金融大模型示范应用,由毫末智行科技有限公司和长城汽车股份有限公司共同开发的自动驾驶大模型DriveGPT示范应用,由北京街远有限公司和北京一轻科技集团有限公司共同开发的面向消费领域的品商大模型示范应用,由北京面壁智能科技有限责任公司和智者四海(北京)技术有限公司共同开发的智能问答大模型示范应用。



输入一张图书馆的图片和一段汽车鸣笛的音频,“紫东太初”就能发现其中的问题,分析出这段音频不太可能出现在图书馆场景中;给出一张救护车图片、一段森林救火的视频和一段警笛音频,它也能准确识别并讲述一段完整的救援故事。

6月16日,在人工智能框架生态峰会2023上,中国科学院自动化研究所所长徐波在发布“紫东太初”全模态大模型时,展示了该模型的新“能耐”。

作为“紫东太初”1.0的升级版,“紫东太初”全模态大模型具备全模态理解能力、生成能力和关联能力,不仅能读懂文字、图像和音频,还能理解视频、3D点云、传感信号,可以结合图像、声音和视频完成场景分析,其理解和生成能力更加接近人类,向通用人工智能迈出关键一步。

全模态升级版更“聪明”

“‘紫东太初’全模态大模型以国产化基础软硬件平台昇腾AI为基础,其算法由中国科学院自动化研究所自主研发,算力则由武汉人工智能计算中心提供支持。”中国科学院自动化所研究员张家俊介绍,无论是算法还是算力亦或是硬件,“紫东太初”都是“中国造”。

“人类学习要受到视觉、触觉、听觉、嗅觉等各种信号的综合影响,每一种不同类型的信号都是一种模态,人类的学习过程是全模态的。同样,机器要想变得‘更聪明’,也需要像人一样,发展贯通图、文、音等多种模态的大模型。”张家俊在谈起开发全模态大模型的原因时解释道。鉴于此,2019年,中国科学院自动化研究所就坚持以“图一言一文”多模态技术为核心,确立多模态大模型布局,整合研究所内图像、文本、语音等研究方向的优势资源开展集团式攻关,于2021年9月成功打造“紫东太初”1.0多模态大模型。“紫东太初”1.0助推人工智能从“一专一能”迈向“多专多能”,向发展通用人工智能迈出了坚实的第一步。

迈入数字经济时代,数据的范畴不断拓宽,不仅包含人类自身产生的语音、图像、文字等数据,还有机器产生的大量结构化和非结构化数据。徐波介绍,针对新需求与新趋势,新升级的“紫东太初”2.0在语音、图像和文本之外,加入了视频、传感信号、3D点云等更多模态,从技术架构上实现了结构化和非结构化数据的全模态开放式接入;突破了能对信息进行充分理解和灵活生成的多模态分组认知编解码技术,以及能融合多个任务的认知增强多模态关联技术等,大模型多模态认知能力大幅提升。

展现广阔产业应用前景

“‘紫东太初’2.0可以像人一样综合运用多种信号进行思考。”据研发团队介绍,它可以理解三维场景、传感信号等数字物联时代的重要信息,完成音乐、图片和视频等数据之间的跨模态对齐,能够处理音乐视频分析、三维导航等多模态关联应用需求,并可实现音乐、视频等多模态内容理解和生成。

“目前,‘紫东太初’大模型已展现出广阔的产业应用前景,在神经外科手术导航、短视频内容审核、法律咨询、医疗多模态鉴别诊断、交通违规图像识别等领域开展了一系列引领性、示范性应用。”徐波说。

比如,在医疗场景下,“紫东太初”大模型与神经外科机器人MicroNeuro相结合,可实现术中实时融合视觉、触觉等多模态信息,协助医生对手术场景进行实时推理判断。

再比如,在法律咨询领域,“紫东太初”大模型仅用0.546秒就可对法律案件进行拆解、提取关键事件,在办公效率方面实现案件分析速度的百倍提升,可有效解决律师整体资源匮乏、服务成本高等问题。

“‘紫东太初’已经开放服务平台,支持各行业根据各自需求‘组装’模块,再输入少量的行业样本数据,就能产出自主可控的行业相关大模型。”张家俊说。

徐波表示,中国科学院自动化研究所未来将以“紫东太初”大模型为基础,持续探索与类脑智能、博弈智能等技术路径的相互融合,最终实现可自主进化的通用人工智能,为促进我国数字经济快速发展贡献力量。

首次举办城市脉搏点亮仪式

全球数字经济大会明日开幕

本报讯(记者 张伟)7月4日,2023全球数字经济大会将在盛大的“数字之夜”仪式中隆重开幕。

此次城市脉搏点亮仪式为首次举办,以数字技术为媒介,选取北京、拉萨和新加坡等城市地标,跨越东盟与喜马拉雅山脉,共同点亮北京城市脉搏。同时,首届中国数字音乐会也将在当晚推出,通过线下互动活动,塑造千年城市的数字形象,提升大会参与感和仪式感。

开幕式当晚,北京市民将在街头感受到数字经济带来的魅力,可以前往首钢园、通州运河、奥运塔、大兴机场和前门大栅栏等城市地标,现场观看炫酷的增强现实(AR)秀。届时,北京城中轴线、东西长安街沿线以及部分市政夜景将全部点亮。

点亮仪式将选取北京城市东西南北中5个方位的5大场景,通过线上和线下联动进行点亮。5个方位分别是西起冬奥首钢园、东至通州运河、北起奥林匹克塔、南至大兴机场以及前

门四点交汇处(中点)。每个地方将演绎本次大会的5个主旨理念:数智、增速、鼎新、赋能和聚合。通过点亮这5个场景,将投射出关键词并展示中心建筑群的演绎,最终呈现出大会的主题。这一点亮仪式不仅辐射整个北京城,分会场还延伸到拉萨以及新加坡。

城市脉搏点亮仪式不仅是对数字经济大会的庆祝,也是展示北京作为数字经济领域的领先城市的机遇。通过点亮城市地标和举办首届中国数字音乐会,大会在北京市民中激发对数字经济的兴趣和热情。同时,通过点亮五大场景和演绎主题,整个北京城被数字经济大会的氛围所包围,将影响力和参与感扩展到世界各地。

2023全球数字经济大会的举办,为北京提供了一个展示其在数字经济领域领先地位的机会。此次盛会将推动本地区的产业发展,也将加强国际间合作与交流,促进数字经济的繁荣与创新。



近日,国产最大直径盾构机——“京华号”盾构机出洞,由中铁十四局承建的北京东六环改造工程西线隧道顺利贯通。这是目前我国最长盾构高速公路隧道,标志着这一国家级创新工程取得重大进展。

新华社记者
鞠焕宗/摄

中国高校最大云上科研智算平台上线

本报讯(记者 张伟)6月27日,国内高校最大的云上科研智算平台CFFF(Computing for the Future at Fudan)在复旦大学正式上线。这台为发现和解决复杂科学问题而建的科研“超级计算机”由复旦大学与阿里云等共同打造,以先进的公共云模式提供超千卡并行智能计算,支持千亿参数的大模型训练——在国内高校中尚属首例,也领先于斯坦福大学等国际知名高校。

延续复旦大学“博学而笃志,切问而近思”的校训,CFFF平台由面向多学科融合创新的AI for Science智能计算集群“切问”一号和面向高精尖研究的专用高性能计算集群“近思”一号两部分组成。

复旦大学教授浩清、人工智能创新与产业研究院院长漆远介绍,基于百G高速数据传输网和阿里云全球领先的大规模异构算力融合调度技术、分级存储技术、AI与大数据一体化技术,部署在复旦校内的“近思”一号和托管在1500公里外阿里云乌兰察布数据中心的“切问”一号连成了一台真正意义上的“超级计算机”,复旦四校区的所有实验设备都能高速接入,做到异构算力统一管理,计算任务统一调度,满足不同应用场景下科学智能研究与应用需求。

位于“东数西算”节点上的阿里云乌兰察布数据中心,通过飞天智算平台的公共

云模式对外提供智能计算服务。

2022年,阿里云发布了飞天智算平台,并启动乌兰察布数据中心等两座超大规模智算中心,为科研、公共服务和企业机构提供强大的智能计算服务,可将计算资源利用率提高3倍以上,AI训练效率提升11倍,推理效率提升6倍。

据介绍,智算中心并不是简单地把配置先进制程芯片的服务器连接起来就能实现高性能算力,其中一个很关键的指标是损耗。智能计算在这个方面难度更大。相比通用型计算,智能计算需要海量数据进行训练,在数据迁移、分布式训练等环节损耗尤其严重。传统智算中心达到一定规模之后,增加算力资源反而会降低算力输出的能力,千卡以上规模的算力输出最低往往仅有40%左右。但阿里云通过体系化的核心技术自研,以飞天智算平台作为输出,大幅降低了智能算力的损耗。

漆远表示,有赖于公共云模式,跑在CFFF平台上的项目可享受到超千卡并行的智能算力,千卡并行的有效算力达到行业领先的92%,可拓展性达到万卡,万卡并行有效算力也可达90%。

同时,公共云更加低碳和绿色。阿里云绿色数据中心技术结合乌兰察布当地天然的气候优势,CFFF平台可实现年平均PUE小于1.2,每年节省总电力超过2000

兆瓦时,年均节碳量达1500吨。

目前,CFFF平台上的第一个科研成果已经诞生。复旦大学人工智能创新与产业研究院李昊团队近期发布了45亿参数量的中短期天气预报大模型,预测效果在公开数据集上首次达到业界公认的ECMWF(欧洲中期天气预报中心)集合平均水平,并将预测速度从原来的小时级缩短到了3秒钟内。

“基于CFFF平台的千卡并行智能计算,这样一个规模的大模型只用一天就完成了训练。传统的计算平台很难做到。”复旦大学人工智能创新与产业研究院研究员李昊说。

这也是CFFF平台上孕育出的第一个大模型。中国科学院院士、复旦大学校长金力表示,复旦希望基于CFFF平台建成一批具有世界级影响力的科学大模型,例如生命科学大模型、材料科学大模型、大气科学大模型、集成电路大模型等。

未来,CFFF平台还将持续扩大其算力规模,并向复旦校外的科研机构、高校、医院、高科技企业等开放。当天,面向全球科研人员的首届世界科学智能大赛宣布正式启动,大赛设置生命科学、量子化学等五大赛道,CFFF平台将为参赛队伍提供免费训练算力,并以更普惠的算力长期支持部分科研项目。