



新看点

“紫东太初”： 从多模态走向全模态

◎本报记者 陆成宽

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“图一音一文”多模态技术为核心，确立多模态大模型布局，整合所内图像、文本、语音等研究方向的优质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、交通违规图像研读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，可在术中实时融合视觉、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。

图说智能

深圳开启无人驾驶试点运营



近日，百度旗下自动驾驶出行服务平台萝卜快跑获得深圳市坪山区授牌，在深圳正式开启L4级无人驾驶商业化试点运营。图为深圳市坪山区的智能网联汽车道路测试与示范应用路段。

新华社记者 梁旭摄

大模型是AI未来发展的重要方向，其研究和应用将逐步成为AI发展的关键方向，并有望形成新一波AI推广浪潮，而AI数据开源将进一步促进大模型的发展。

林咏华

北京智源人工智能研究院副院长兼总工程师

数据开源为AI发展“推波助澜”

◎实习记者 裴宸纬

在6月9日—10日举行的2023北京智源大会上，“AI数据开源”引发广泛关注。AI数据为什么要开源？AI数据开源面临哪些挑战？它会是未来AI发展的重要趋势吗？科技日报记者带着这些问题采访了相关专家。

AI数据开源意义重大

有专家认为，AI数据开源对深度学习模型的发展意义重大。由于训练AI大模型需要大量资源，所以预计“赢家通吃”类AI系统的开发和管理将首先由少部分闭源实体所主导。

但遗憾的是，这种资源限制导致研究人员、非营利组织和初创公司等小规模实体因无法承担高昂的成本，几乎不可能从零开始训练自己的AI大模型。

以对话类模型为例，目前国内外众多已经开源的对话模型，其实都是基于语言基础大模型，再利用少量指令微调数据进行训练所得。

如果开源AI大模型的数据在质量上具有足够的竞争力，深度学习模型的规模化训练和运行成本将大幅降低。

北京智源人工智能研究院（以下简称智源）副院长兼总工程师林咏华对记者表示，大模型是AI未来发展的重要方向，其研究和应用将逐步成为AI发展的关键方向，并有望形成新一波AI推广浪潮，而AI数据开源将进一步促进大模型的发展。

深度学习需要大量的标注数据进行模型训练。在林咏华看来，过去10年，深度学习技术快速发展的重要原因，就是许多志愿者团体、国外科研团队一直在积极地收集、整理并开源用于深度学习的训练数据集。“当前AI大模型训练对数据量的需求，比之前的深度学习小模型对数据量的需求有了百倍，甚至千倍的提升。所以，尤其在过去一年，数据开源的问题日益受到广泛关注。”林咏华说。

背后挑战不容忽视

开源自然会为AI发展带来诸多好处，但其背后的挑战

也不容忽视。其中之一，便是开源安全与合规挑战。林咏华认为，对传统的商业软件而言，开源中的安全、合规、许可证和代码质量风险等是使用开源组件必须面临的挑战。然而在AI大模型时代，更大的挑战则在开源数据集方面。

因此，AI数据开源应在协议许可的范围内进行。“用于AI大模型训练的开源数据必须是合法地从公开或可公开获得的资源中收集的数据。人们可以在开源协议允许的范围内，以AI大模型训练、AI算法开发为目的，对数据进行访问、修改和使用。部分数据可能要求使用过程中遵守更严格的协议。”林咏华表示。

此外，今天的基础AI大模型不只具备理解能力，还具有生成能力，它能够对外进行认知输出、价值观输出等，可能给社会带来巨大影响。“我们在训练基础大模型的时候，所使用的预训练数据会对AI生成内容质量起到很大程度的决定性作用。因此，开源数据的质量十分重要。”

林咏华指出，由于高质量的数据（如文章、图片、视频等）通常有版权，由于版权或商业因素导致的闭源以及数据孤岛等挑战会制约AI的发展，所以需要多方推动构建更多高质量的开源数据集，尤其是用于训练基础AI大模型的开源数据集。

LF AI & DATA基金会董事主席堵俊平对此也深有感触：“AI大模型就像一个贪吃的‘怪兽’，始终需要研究人员投喂更多的、质量更好的数据。”他说，当前数据几乎都是从“在网络上主动收集”“从第三方购买”“利用公开数据集”这三个渠道得来。在堵俊平看来，从第一个渠道得到的数据局限性较强，由于版权问题，很多公司只能从其私域获得数据；从第二个渠道获取的数据面临数据定价、数据质量等问题；而从第三个渠道获取的数据往往只能作为研究使用，在商用或者其他方面有很多限制。

开源渐成AI发展重要趋势

记者了解到，智源对2023年1月到5月底发布的、具有影响力的语言模型进行过统计。统计结果表明，国外发布的开源语言模型有39个，国内发布的开源语言模型有11个。

“开源是推动AI技术进步的重要力量，AI开源开放生

态及平台建设也日益受到重视。开源开放毫无疑问已经成为重要的AI发展趋势之一。”林咏华表示，“开源能够促进AI大模型科研创新，推动和降低AI大模型落地乃至整个AI产业落地的门槛。”

然而，通往开源的道路并非一帆风顺，在数据之外，算力也是开源路上的一只“拦路虎”。AI大模型训练依赖庞大的数据、算力。训练参数量级的增长使得算力需求也随之增长，算力集群正变得愈发庞大。

然而算力成本却是小型开发者的“不可承受之重”。拿到AI大模型开源数据后，往往需要对其进行微调和二次开发。但现实的情况是，对一些小型开发者来说，仅仅是做推理都很困难，就更别提对AI大模型做微调、二次开发。以ChatGPT为例，仅就算力而言，Open AI为了训练它，就构建了由近3万张英伟达V100显卡组成的庞大算力集群。有消息称，Open AI公司发布的新一代语言模型GPT-4甚至达到了100万亿的参数规模，其对应的算力需求同比大幅增加。

目前，有一些研究机构希望用技术的革新抵消巨大的算力成本。最直接的手段是通过训练技术的革新加快AI大模型推理速度，降低算力成本，减少能耗，以此来提高AI大模型的易用性，让开源数据更好地发挥价值，但这只能从工程上对算力资源的约束起到缓解作用，并非终极方案。

有业内专家表示，解决算力问题最终还是要回到AI大模型自身寻找突破口，一个十分被看好的方向便是稀疏大模型。稀疏大模型的特点是容量很大，但只有用于给定任务、样本或标记时，模型的部分功能才会被激活。也就是说，这种稀疏大模型的动态结构能够让AI大模型在参数量上再跃升几个层级，同时又不必付出巨大的算力代价，一举两得。

此外，开源社区的作用同样不容忽视。开源社区是推动开源发展的重要基石，开源的最初发源地，就是来自于社区开发者的贡献。“Linux系统的成功很大程度上得益于开源社区。30多年来，Linux系统发展成为拥有海量全球用户的操作系统，其成功以及长久不衰的秘诀就是开源，尤其是内核社区成千上万开发者的贡献。”林咏华举例说。

“开源开放可以使得我们站在前人的肩膀上前行。”林咏华总结道，“这些年AI领域取得的成果大多受益于开源，如果没有开源，AI不会发展到今天。”

行业大模型落地将加速产业升级

◎本报记者 操秀英

6月19日，腾讯云行业大模型及智能应用技术峰会在北京举行，腾讯云首次公布了其行业大模型的最新进展，并与来自教育、传媒、政务等领域的生态伙伴，分享了对行业大模型的思考。

产业应用场景成为大模型最佳“练兵场”

对于大模型的产业价值，中国科技新闻学会党委副书记、副理事长兼秘书长陈锐表示，大模型就是要立足行业应用，立足智慧民生。

当前，我国产业正面临智能化转型，人工智能与国内各行业深度融合，成为驱动产业转型升级的重要力量。腾讯公司董事会主席兼首席执行官马化腾也表示：“我们最开始以为人工智能是互联网十年不遇的机会，但是越想越觉得这是几百年不遇的、类似工业革命一样的机遇。”

以制造领域为例，据不完全统计，中国已建成2500多个数字化车间和智能工厂，带动汽车、家电等离散型工业的产品迭代速度和交付能力明显增强，钢铁、石化等流程型工业的本质安全和绿色发展水平大幅提升。

也有专家认为，“十三五”以来，我国在

智能制造领域取得了长足进步，但新一代信息技术和先进制造技术的融合程度还不够，应逐步尝试基于新一代人工智能的智能制造。

因此，随着人工智能技术的不断升级，推动大模型产业化应用，部署建立行业大模型，是我国产业升级的重要举措。

当前，我国众多产业对于高质量发展的需求，将为大模型的落地应用提供场景支撑。可以说，产业应用的实际场景已经成为了大模型的最佳“练兵场”。

平台型互联网云厂商助推行业大模型发展

推动人工智能发展，要坚持“基础研究和产业应用”两条腿走路。科技部此前已经表态，将把人工智能作为战略性新兴产业和新增长引擎，一方面要推动构建开放协同的人工智能创新体系，加快基础理论研究和重大技术攻关，另一方面要推动人工智能与经济社会深度融合，在重大应用场景中锤炼技术，升级迭代，培育市场。

有专家认为，在人工智能基础研究方面，我国虽起步较晚，但是近15年的发展速度全球首屈一指。与此同时，在人工智能产业应用方面，我国已进入全球第一梯队。根据中国信息通信研究院的统计，2022年我国人工智能核心产业规模达5080亿元，同比增长18%。有专家预测，

随着人工智能技术的不断升级，推动大模型产业化应用、部署建立行业大模型是我国产业升级的重要举措。我国众多产业对于高质量发展的需求，将为大模型的落地应用提供场景支撑。

2030年我国人工智能产业规模将达到1万亿元。

广阔的市场机遇，给行业大模型带来了巨大的发展空间，这也对行业大模型产业链的生态协同，提出了新的要求，同时要避免重复建设、资源浪费、数据孤岛等情况的出现。

去年年底的中央经济工作会议指出，支持平台科技企业在引领科技创新中大显身手。因此，平台科技企业应发挥自身规模优势，成为数字关键核心技术自主创新的主力，尤其是产业升级对于行业大模型的迫切需求，平台型互联网云厂商大有可为，可以为行业大模型建设提供助力。