

人工智能领域2023年是属于“大模型”的一年 “百模大战”抢滩智慧未来



图 IC

本报记者 郜阳

如果你还没有听说过“大模型”，那么你的生活距离科技可能有点远了。如果说去年人工智能领域的关键词是“元宇宙”的话，那么2023年无疑是属于“大模型”的一年。迄今，国内的科技公司、高校、研究机构已经发布了数十款人工智能大模型产品，“百模大战”一触即发。在即将于上海举行的2023世界人工智能大会上，也将汇聚国内外大模型明星团队。其中，百度文心、阿里通义、华为盘古、讯飞星火、商汤日日新等30余个大型模型悉数登场。大模型语料联盟、测评平台、测试标准及AIGC白皮书等一批重磅成果将在大会期间集中启动或发布。逐鹿“大模型”，会如何左右人工智能的发展，又最终会给生活带来什么样的影响？本届大会，或许会给出一些答案。

『大力出奇迹』超普通用户预期

很多人或许没有在意，在去年的世界人工智能大会上，大模型已经悄然来到了我们身边——中国科学院自动化研究所和华为联合开发的全球首个三模态大模型“紫东太初”斩获大会最高奖项SAIL奖，在工业质检、影视创作和手语教学等领域展现不错的能力。

但你要问普通人对“大模型”的认识，更多人会给出ChatGPT的答案。的确，自从发布之后，它在短短两个月内就拿下了活跃用户1亿的成绩，远远甩开了其他应用。ChatGPT写诗、写小说、敲代码，几乎无所不能，掀起一场人工智能风暴。ChatGPT之所以火出圈，特别之处在于，它被普通用户观察到，并超过了人们对于人工智能的期待。

据业内人士分析，以ChatGPT为代表的大模型目前在多个方面有价值：帮用户生成一段内容作为初稿，用户再去修改；把信息有条有理地呈现出来，给用户提供启发和灵感；当然，还有娱乐的功能。

首届世界人工智能大会或许是一道分水岭。在它启幕前，小模型是人工智能理解人类语言的主流技术，最为人们熟知的可能就是下围棋了——标注好数据，进入模型训练。不过也好理解，小模型无法用于其他任务，数据标注成本也极高。当时的人工智能行业，在机器学习等诸多技术领域已然进入瓶颈。

后来，谷歌发表论文，引入自注意力机制学习文本，神经网络需要学会自动判断哪些词语对理解句意最有帮助，而不是“全文死记硬背”，不再像过去一样需要大量精标样本，有人戏称，大模型某种意义上是“大力出奇迹”的“暴力美学”——充分利用网上海量的未标注数据，模型的规模和能力得到显著提高，加之大模型具有非常强的通用能力，只需经过少量参数微调，就可以用于不同任务。

科大讯飞董事长刘庆峰告诉记者，认知智能大模型智慧涌现，有望改变现有的信息分发获取方式、革新内容生产模式、全面升级人机交互方式以及加速“AI for Science”的发展，将有可能给整个社会带来颠覆式影响。

2 // 考量因素不再“简单粗暴”

可以说，人工智能大模型是在一夜之间热得发烫的。

仅仅是去年，人工智能还因商业化困局而被外界诟病，而ChatGPT好似搅动大模型江湖的鲶鱼，一时间，各类选手纷纷“下场”——《中国人工智能大模型地图研究报告》显示，截至6月，我国10亿级参数规模以上大模型已发布79个，现已在大模型方面建立起涵盖理论方法和软硬件技术的体系化研发能力，形成紧跟世界前沿的大模型技术群。

移动互联网时代的第一场硝烟弥漫的“千团大战”，或许已经渐渐被人淡忘，而人工智能时代的“百模大战”则将掀起又一次的持久战：百度在今年3月16日发布了“文

心一言”；阿里的“通义千问”从4月7日开始企业内测；商汤科技在4月10日的技术交流会上推出“日日新”大模型体系；科大讯飞在5月6日举行星火认知大模型成果发布会……

“百模大战”之际，各家的大模型能力究竟几何，是最令外界好奇的。以往，“大厂选手”会“简单粗暴”地比拼参数量，可这个标准早已不能当作衡量大模型能力的唯一参考了。在中国工程院院士贺铨看来，评价大模型水平应该是多维度的，全面性、合理性、使用便捷性、响应速度、成本、能效等。

尽管如何评价大模型的实力还缺乏统一标准，但大模型的能力已然可以通过学术层面的评估以见一斑——面向人类设计的

高难度、综合性考试被越来越多地引入对大模型的智能水平进行评测。

在今年高考首日，上海人工智能实验室发布千亿级参数大语言模型“书生·浦语”，联合团队选取了20余项评测对其进行检验，结果显示：“书生·浦语”在多项中文考试中取得超越ChatGPT的成绩，其中就包括中国高考各科目的数据集(GaoKao)。

上月下旬，商汤科技公布了自研中文语言大模型“商量 2.0”在MMLU、AGIEval、C-Eval三个权威大语言模型评测基准的成绩。根据评测结果，“商量 2.0”在三个测试集中表现均领先ChatGPT，部分已十分接近GPT-4的水平。

“今天的大模型和昨天的大模型，是完全不同的两个。”这是记者在走访研发大模型的科技企业时，听到最多的一句话。

3 // 开放平台带来更繁荣生态

当下，生成式人工智能行业鲜花着锦，诸多大模型争相问世，但其应用仍面临能否深入行业场景的问题。

星环科技市场部总经理谭彬介绍，今年是星环科技第六次参加世界人工智能大会，将在大会期间全球首发金融领域大模型星环无涯产品，“在金融量化领域，‘无涯’可回答研报分析，能对个股、债券、基金等各类市场事件进行复盘和推演，有力辅助分析师、研究员和投资经理的日常工作。”他透露，企业也会展示大数据分析大模型“求索”，“用户通过自然语言即可生成可执行SQL语句，快速获取数据库查询结果，大幅度降低数据科学家门槛。”

在科大讯飞星火认知大模型问世1个多月后，其V1.5版本发布，除了开放式知识问答取得突破，多轮对话、逻辑和数学能力再升级之外，其应用落地更引人关注。

记者关注到，讯飞星火认知大模型V1.5

更进一步拓宽了对多行业领域的赋能。例如在医疗领域，讯飞诊后康复管理平台已落地医院。科大讯飞总裁吴晓如介绍，在患者出院康复的过程中，需要个性化的康复服务指导，但目前诊后延续性服务普遍缺失。而基于迭代优化的大模型技术，诊后康复管理平台可将专业的诊后管理和康复指导延伸到院外——根据患者健康画像自动分析，平台可为患者智能生成个性化康复计划，包括重点关注、用药指导、康复运动、出院随访、健康知识、患者咨询等，并督促患者按计划执行。

“讯飞人工智能开放平台目前的开发者团队超过400万家，这一次通用人工智能所带来的生态将更加繁荣。”刘庆峰表示。

百度创始人、董事长兼首席执行官李彦宏在前不久举行的中关村论坛上预测，今后所有的应用都将基于大模型来开发，每一个行业都会有属于自己的大模型，大模型会深

度融合到实体经济中去。

当然，很多人提出焦虑：随着大模型的发展，未来会不会替代人类的工作？

这个问题，或许首先要问问科学家们，从蛋白质结构预测到气候系统建模，从引力波探测到理解宇宙，在很多领域，人工智能大模型正以较以往数倍甚至数万倍的效率，改变多学科的研究。在晶泰科技创始人、董事长温书豪看来，伟大的数学家会问为什么1+1=2，然而AI不会。AI会成为超级工具，会帮助人类记住无法记住的知识、无法计算的复杂方程，但是，深度的思考和提问的能力，人依然是独特而不可替代的。

记者也在一款语言大模型里输入了这个问题，得到了这样的回答：大模型在某些领域可以提高工作效率，但它们不会完全取代人的工作。人工智能和机器学习技术的发展使得计算机能够处理大量数据、执行复杂任务并提供智能决策支持；然而，人类在很多方面仍然具有独特的优势，如创造力、情感智慧、沟通能力和适应性等。

4 // 中国必须做自己的大模型

显然，大模型并非一朝一夕就能够炼造的，现如今“百模大战”战况不断，本质还是过去几年技术积累的产物。不少专家指出，大模型相关新研究、新产品竞相涌现，通用人工智能正成为各国重点布局的关键赛道。

“抢抓通用人工智能的发展机遇有几个基本要素。”科大讯飞董事长刘庆峰认为，第一，必须要在自主可控的平台上；第二，必须要同时做中文和英文，不只学习中国的“智慧”，还要向世界学习；第三，在“硬碰硬”的科技对比上，不仅要学习，还要想办法赶超。

“中国必须做自己的大模型”，这已经成为共识。为推动人工智能创新发展，北京、

上海、广东深圳等多地纷纷出台支持举措，集聚人工智能创新资源，抢抓大模型机遇，开展大模型创新算法及关键技术研究，打造人工智能创新高地。

日前公布的《上海市加大力度支持民间投资发展若干政策措施》提出，充分发挥人工智能创新发展专项等引导作用，支持民营企业广泛参与数据、算力等人工智能基础设施建设。

徐汇区副区长俞林伟在2023世界人工智能大会新闻发布会上透露，徐汇在游戏、医疗、金融、零售等领域，已经汇聚了一大批垂类大模型企业，将全力支持他们自研技术创新，加快打

造有机循环产业发展生态。

“大模型已经引发办公软件、内容创作、交互终端等领域的‘新工业革命’，业界也有共识，‘所有行业都值得用大模型重做一遍’。”俞林伟表示，“徐汇将直面企业需求和产业关切，通过率先开放应用场景等政策，引导行业内领军企业主动深化改革、提升能级，加快拥抱这一轮人工智能的发展。”

同样需要正视的是，未来，要对大模型设置安全“围栏”。业内人士建议，合理利用法规应对人工智能带来的风险与冲击，形成安全“护栏”，并在其引导下营造更有活力的发展环境。