

前沿周刊 / 科技

扬帆逐浪:迈向人工智能高地的浙江脚步

—AI浪潮系列观察之三

■ 本报记者 何冬健

在新一轮人工智能的浪潮中,浙江的存在感格外强烈——

深度求索(DeepSeek)和通义千问(Owen)大模型综合性能处于全球第一梯队;“三体计算星座”成为我国首个进入组网阶段的整轨互联太空计算星座;“魔搭社区”吸引全球超2000万开发者,成为国内最大的开源社区;宇树科技、群核科技等杭州“六小龙”引起全球关注……

根据省委、省政府部署,浙江加快打造人工智能创新发展高地。杭州在政府工作报告中提出全力建设“人工智能创新发展第一城”。人工智能就是赋能浙江未来发展的核心变量,已成为全省上下的共识。

当全球AI竞争逐渐进入深水区,浙江迈出的每一个探索脚步,都在为中国人工智能持续创新发展提供实践和思考的样本。

从感受危机到快速突围

“原来一年前我们以为上了船,后来发现那个船漏水了。现在感觉站了上去,还坐不下去。还是希望船能快一点。”前不久,腾讯创始人马化腾在被问及腾讯AI是否落后时这样表示。

这番话或许说出了一大批中国AI从业者的集体感受。过去几年,所有人都在找船、上船、修船。这种危机感,在浙江,有人更早感受到,也更早布局。

“在ChatGPT诞生以前,中国AI人的心境大致是从容的。我们在追赶,差距在缩小,时间站在我们这边。”浙江省人工智能学会一位专家回忆。直到2022年11月30日的夜里,ChatGPT悄无声息上线。没有发布会,没有预热,短短两个月,月活破亿。阿里云园区的研究员彻夜未眠,他们比绝大多数人都更早嗅到了危险。那个夜晚,“船漏水了”成为整个行业的共同体验。

船漏了,需要一群人不分昼夜地干,把木板重新钉上。时间来到去年4月28日,杭州三墩灯彩街1008号的阿里云园区,又是一个不眠夜。Owen3发布后,很快登顶全球开源模型榜首。而在此前3个月,深度求索已引爆全球,刚发布,便超越ChatGPT登上全球140多个应用商店榜首。

3年不到的时间,从追赶到跻身世界前列,这群人做对了什么?

早在2020年,全球AI圈还在为图像识别和推荐算法竞争,原达摩院智能计算实验室负责人周靖人就拉起了一支十几个人的小团队,筹划一个名为“M6”的超大参数规模AI预训练模型。他们创造性地提出MoE(混合专家)架构——让一个大模型根据不同任务,调用内部多个专家。通过自研的Whale分布式框架和专家并行策略,只用480张卡就完成了万亿参数模型的训练。和英伟达用3072块A100训练万亿参数、谷歌用2048块TPU训练1.6万亿参数相比,能耗降低了超过八成,效率提升了近11倍。MoE这个架构,今天成了行业标配。但当年十几个人的团队敢这么干,需要一点“不信邪”的劲头。

“中国的研究人员有一个特点——愿意做那些‘不够光鲜’的工作。”今年5月初,美国AI研究员内森·兰伯特在密集访问了7家中国AI实验室后,给出这样一个结论。大模型能力提升,很多时候并不来自论文里那些惊艳的创新,而来自无数细节的持续优化。数据清洗、参数调校……这些工作枯燥、耗时,做好了也未必能发一篇像样的论文。但在浙江,有大量的AI研究人员愿意扎根在这些基础工作上,只为了最终模型效果再好一点点。

除了这种“坐冷板凳”的耐心,还有另一重优势:年轻。过去几年,大模型技术路线变化一轮接一轮。从MoE到强化学习,再到Agent智能体,几乎每隔一段时间,行业范式就会发生一次翻转。那些经历过上一轮AI周期的老兵,有时难免被经验束缚。年轻一代研究人员,没有太强的历史路径依赖,更容易“轻装上阵”,快速接受新的技术方向。

“我们没包袱。”一位接受采访的95后研究人员说,“既然这条路是对的,就要走下去。”

“浙江三剑客”同台起舞

今年4月下旬,深度求索在“迟到”半年后发布V4版本大模型,好评如潮,中外热搜上了一遍。这家以开源模型名声大噪的创业公司有个外号——“真正的Open AI”,业内都佩服其“品味



2025年世界互联网大会乌镇峰会,“六小龙”乌镇对话。

潮新闻记者 李翔 摄



阿里通义千问大模型家族。受访者供图



中国地质科学院地质研究所副所长王瑶(右一)在2024年美国地球物理学联合会向公众展示OnePetrology(岩浆岩数据库)与GeoGPT合作成果。受访者供图



在浙江大学杭州国际科创中心人形机器人创新研究院,研发团队在训练机器人博尔特。新华社发

好”。

回看2023年7月,初创的深度求索还毫无名气,充满理想主义色彩。“后边可以把我们的训练结果大部分公开共享……而不是技术只掌握在一部分人和公司手中,形成垄断。”公司创始人梁文锋在接受媒体采访时曾这样说。他们从一开始就走了一条不同的路——极致性价比:用尽可能少的算力资源,做出更实用的模型。当DeepSeek-R1以低得多的成本逼近世界顶尖模型性能时,整个硅谷都感受到了震动。

有专家告诉记者,深度求索的成功,让业内开始认识到,中国头部实验室与世界前沿实验室之间的差距,远没有外界想象得那么大。

如果说深度求索代表了初创公司以极致创新撕开一道口子,那么阿里巴巴则充分展现了龙头企业进行系统性布局的定力与实力。

2008年9月,阿里就确定云计算和大数据战略。宽容和开放的环境,十几年的投资,最终让阿里得以在大模型时代手握全球前四的云计算公司这张牌。“云+AI”的组合,加上阿里的Token生态,长久、稳健地推进大模型朝着全场景、全模态、全尺寸方向发展。

在之江实验室,AI研究者在干一件“有点疯狂”的事。他们想把基因序列、蛋白质结构、天文数据、材料分子结构——这些完全不同学科的科学知识,统统变成一种统一的数据语言,放进“021”科学基础模型。

为什么叫“021”?在英文里就是“0到1”的谐音,意味着之江实验室在做一件“从0到1”的创新。这种原创性追求,正是之江实验室作为“三剑客”中新型研发机构的根本所在,它的使命是探索长远看极为重要的方向。



杭州湾具身智能创新中心开展机器人作业模型训练,为具身智能机器人进入真实生产线积累经验。新华社发

之江实验室有这个底气。他们选择了一条最难的路——从零开始构建一个为科学而生的基础模型。这需要从“地基”开始造房子。

研发团队充分发挥MoE与科学知识结构高度同构性的优势,构建“科学数据统一表征+MoE”融合模型架构,探索将科学数据和文本语料编码到统一的高维空间,让模型能够识别、处理科学数据,认识并解决复杂科学问题。在生命科学领域,之江实验室与华大生命科学研究院的专家奋战6个月,打造了全球首个百亿级人类基因组基础模型Ge-

nos,其在临床致病性突变解读方面准确率达92%,结合之江实验室“021”科学基础模型后更提升至98.3%,为临床诊断提供了全新的高效工具。

大企业、初创公司、新型研发机构,三种打法,在浙江形成了一种极为罕见的互补格局。

让AI大河充盈源头活水

登上AI这艘船,行至中流,风景愈发壮阔,风浪也更大。

浙江省科技信息研究院院长俞锋华向记者细数浙江AI创新发展的优势和短板。当前,浙江在人工智能等领域拥有之江实验室等高能级科创平台,以阿里云、海康威视、新华三等为代表的全球智能企业,深度求索、群核科技等大模型主力军,构建了“国际领跑、国际并跑与国内领先”协同发展的技术和产业体系。

特别是人才网络,堪称浙江的特殊优势。浙江大学每年向省内输送数千名计算机方向的人才,阿里巴巴多年来培养了大批AI工程师和产品经理,这些人积累经验后,继续在杭州、在浙江创业。回看2020年阿里云M6项目的专家团队,有不少现在业界的知名人物。比如清华大学教授唐杰后来创立了智谱AI,而他的学生杨植麟创立了“月之暗面”,做出了Kimi大模型。一个十几人的小团队,生长出了中国AI领域的多家头部公司。浙江的AI土壤,从那时就开始积蓄力量。

浙江的短板也很明显。国产AI芯片性能与生态建设仍需加强,高性能科学仿真计算平台仍未完全实现国产自主;车规芯片、激光雷达、高端工控软件等关键环节的供应链稳定性面临考验;实用性量子优越性算法尚未成熟,复杂工况下量子传感器探测潜力受到限制等。即便是在发展速度最快的杭州,和

链接

五大路径,解码浙江打造AI高地

当前,浙江正在打造人工智能、生命健康、新材料新能源三大科创高地。前不久,省科信院联合之江实验室发布《2025年度人工智能科创高地发展报告》,聚焦云计算与大数据技术、智能计算与系统、微电子与光电子、智能控制与先进制造、量子科技五大战略领域。

围绕这五大领域,报告明确了五大重点举措:

一是加强重点领域关键核心技术攻关。重点突破云智算、人工智能数据基、智能机器人、科学智能计算、类脑计算等。加快通义千问、深度求索等全模态大模型技术攻关与迭代。

二是加快培育人工智能战略科技力量。重点推动3家省实验室、5家省技术创新中心提升科创平台策源能力。加快省CMOS集成电路技术创新中心等概念验证中心和中试平台建设。面向智能网联汽车、智能体等重点领域,加大模

型应用成效和控制安全验证。

三是打造人工智能企业梯次成长新生态。推行“四题一评”协同攻关机制,支持人工智能企业牵头承担自主框架构建、高性能AI芯片开发等科技计划项目。完善科技领军企业培育库建设,支持符合条件的人工智能领域企业入库培育。

四是建设人工智能赋能科学研究高地。依托海纳数据枢纽,打造“浙江省科学数据汇交与服务平台”,建设融算力、数据、模型为一体的开放创新基础设施。联合重点科创平台共建高质量科学数据集。

五是建立实战导向的AI人才引育体系。支持人工智能芯片重点企业引进顶尖人才,放宽海外学习(工作)经历条件限制。建好省人工智能学院等平台,将学历教育融合到研发实训过程。

(本报记者 何冬健 整理)