

# 张勇的“云”端战事



## 赶时髦

阿里架构调整后,张勇带着通义千问来到阿里云峰会现场。

4月7日,阿里云自研大模型“通义千问”面向企业用户内测。4月11日,阿里云宣布未来阿里所有产品将接入“通义千问”大模型全面改造,并向客户开放通义千问的能力。先推出产品,再自我改造、开放能力,微软、谷歌、百度、阿里云走的是同一个路子,产品间的功能性差异不大,AI生成的满意度见仁见智。

作为首批内测用户,北京商报记者对比通义千问和文心一言发现通义千问在交互主页面下方设置了小应用入口百宝袋,将通义千问的能力归纳为效率类、生活类、娱乐类三种,文心一言的交互界面相对简单。两者都可以进行多轮对话、内容创作等,但通义千问目前不支持画图、语音、视频等多模态生成。

以写诗功能为例,北京商报记者分别要求通义千问和文心一言以自己的产品名写一首藏头诗。通义千问的作品是:通天达地任我行,义薄云天我自守。千山万水行远路,问天不语心自知,文心一言写下:文章如日月,心气横骨鲠。一代帝王师,言论安邦国。

“坦诚地讲,GPT在各个方面处在领先的地位,我们需要取长补短,希望在下一个阶段能贡献一些技术创新。”阿里云智能CTO周靖人向北京商报记者承认有差距,同时强调“通义千问不是为了针对ChatGPT做的,而是阿里云在预训练大模型探索的一个方面”。

张勇来了。4月11日,同行、客户、媒体、合作伙伴盯着独立后的阿里云智能和新任CEO张勇的一举一动,听他回忆与阿里云的合作,听他解读给阿里云开出的药方:“未来阿里所有产品将接入‘通义千问’”“阿里云希望帮助更多企业用上大模型”……

在张勇之前,阿里云经历三代掌门人。负责从0到1的王坚、从技术跨越到商业的胡晓明和平衡技术和盈利的张建锋,没让阿里云掉下国内云计算老大的位子,也眼看着竞争对手陆续登场。曾保持谨慎的腾讯CEO马化腾和百度CEO李彦宏回过神来,暴雨中奔跑的华为不成功便成仁,云计算不只是阿里未来的第二增长引擎,生成式AI正经历“诸神之战”。

张勇接手阿里云时,营收增幅放缓也被认为是挑战之一。2022年12月张勇接管阿里云时,文渊智库创始人王超曾向北京商报记者直言,“这次调整表达了阿里集团对阿里云过去两年发展的不满。阿里云智能未来在阿里集团的地位将越来越重要,但我猜测张勇兼任可能是过渡。”阿里被“拆分”后,王超改了主意,“看来张勇要常驻阿里云了”。

2023年3月28日,张勇在阿里集团之下设立了六大业务集团和多家业务公司,自己兼任阿里云智能CEO,阿里云的“绑定”越来越紧。在阿里云峰会现场,张勇也提到了自己身份的变化。回忆起当年作为阿里云客户时的情景,2019年、2010年让他印象深刻。当时阿里已经经历过“双11”带来的流量峰值挑战,云计算如何能够支撑着这样的“洪峰”,在成本上又是个可持续方案?阿里集团和阿里云没有信心,于是一个如今听起来非常可笑的可笑的备用方案被提出,“实在没办法就把淘宝页面切成黑白的”。

现在阿里云再也不用为一个流量洪峰支撑系统而发愁。“非常幸运,阿里云赶上了中国数字产业发展周期和数字中国战略,让我们有机会了抓住移动互联网的机遇、千行百业数字化的机会。”张勇直言。

AI大模型加速了智能时代的到来,在张勇看来,所有企业在智能化时代都将站在同一条起跑线上,互联网公司和非互联网公司不存在先进或者落后的区别,传统企业有弯道超车的机会,所有产品都值得用大模型重做一遍,未来将看到应用爆发、效率革命,AI将成为各行各业新型生产力。

## 零门槛

竞争对手也看到了大模型时代的机会,甚至想在这一局打掉阿里的先发优势,同类产品包括但不限于百度的文心一言、腾讯的“混元大模型”、华为的盘古大模型。创新工场CEO李开复、美团元老王慧文、搜狗创始人王小川……每个人兜里都能掏出一把AI好牌。

微妙的时间差被外界察觉,谁都不想做攻城者。2022年9月,李彦宏公开表示,人工智能发展在“技术层面和商业应用层面,都有方向性改变”。2022年11月,阿里云在云栖大会上推出ModelScope模型即服务(MaaS)平台。

“阿里通过阿里云实现对生态伙伴的云端赋能,现在这个载体转向了大模型。”陈端解释。

未来十年,阿里云将为AI时代提供好算力基础和大型模型能力,全面服务智能化进程。为此,阿里云要做好核心两件事:让算力更普惠、让AI更普及,张勇给阿里云定了调。

具体包括,2023年阿里云将推出一款更接近云计算终极形态的算力产品——通用Universal实例,这款产品进一步屏蔽了传统IT的硬件参数,让数据中心真正成为一台超级计算机,为中小企业和开发者提供普惠的算力。通用Universal实例的价格,对比上一代实例最高可下降40%。

除了帮助用户优化成本外,阿里云还向开发者们抛出橄榄枝,推出“飞天免费试用计划”,面向1000万云上开发者提供包括ECS、数据库PolarDB、机器学习PAI等在内50款云产品的免费试用。同时还提供完备的产品文档、一键部署的技术解决方案、1000多门免费课程,让开发者“零门槛”体验云服务。

听到周靖人的降价通知,坐在北京商报记者身边的一位阿里云客户挺直了身子。

北京商报记者 魏蔚/文 企业供图

## 刚开始

我们为什么计算?

在张勇上场前,会场大屏幕滚动播放着王坚出镜的一则宣传片,他以电气化时代讲起,回答了这个问题,“现代计算机的发明满足了我们个人的计算需求,直到云计算的出现,你才真正拥有了在任何时候,在地球的任何角落获取计算的能力。云计算正在重新定义今天的世界,而这一切才刚刚开始”。

大模型的竞争也刚刚开始。目前,钉钉、天猫精灵已在测试接入大模型,接入通义千问后的钉钉可实现近10项新AI功能,比如自动总结未读群聊信息中的要点等。天猫精灵也会有新的科技和狠活,用户可以在跑步时要求天猫精灵“合成1小时歌单,50%穿插摇滚风格的歌曲”。

“通义千问有明显的阿里色彩,它的功能跟阿里之前的积累深度嵌套,偏重于面向C(用户)端的电商、办公、服务的流程优化、场景优化和体验优化,给用户带来一个更好的消费附加价值和消费体验。”中央财经大学数字经济融合创新发展中心主任陈端向北京商报记者表示。

“当应用在实际场景时,企业可能有特殊的需求,也积累了一些行业经验,如何把这些知识体系融入到通义大模型里,这是企业专属大模型要解决的问题。”周靖人解释,“不是所有企业都需要从头开始训练。”

云计算是AI大模型运行的基础,“就像ChatGPT的成功受益于微软的云计算服务Azure,阿里的通义千问也受益于阿里云。未来持续应用AI需要算力开销,对阿里云的业务拓展是一个积极助推”。浙江大学国际联合商学院数字经济与金融创新研究中心联席主任盘和林向北京商报记者表示。

这种助推正是阿里云急需的。根据Canalys最新数据,2022年四季度,国内整体云计算基础设施支出(IaaS基础设施即服务+PaaS平台即服务)以增长4%收尾,全年支出303亿美元。对比前三年,超过30%的年增长率跌至10%。

阿里云的营收弧线和大盘一致。2022年四季度阿里云营收同比增速3%,和2018年四季度超80%的增幅形成强烈对比。

## 幸运儿

张勇接手阿里云时,营收增幅放缓也被认为是挑战之一。

2022年12月张勇接管阿里云时,文渊智库创始人王超曾向北京商报记者直言,“这次调整表达了阿里集团对阿里云过去两年发展的不满。阿里云智能未来在阿里集团的地位将越来越重要,但我猜测张勇兼任可能是过渡。”阿里被“拆分”后,王超改了主意,“看来张勇要常驻阿里云了”。

2023年3月28日,张勇在阿里集团之下设立了六大业务集团和多家业务公司,自己兼任阿里云智能CEO,阿里云的“绑定”越来越紧。

在阿里云峰会现场,张勇也提到了自己身份的变化。回忆起当年作为阿里云客户时的情景,2019年、2010年让他印象深刻。当时阿里已经经历过“双11”带来的流量峰值挑战,云计算如何能够支撑着这样的“洪峰”,在成本上又是个可持续方案?阿里集团和阿里云没有信心,于是一个如今听起来非常可笑的可笑的备用方案被提出,“实在没办法就把淘宝页面切成黑白的”。

现在阿里云再也不用为一个流量洪峰支撑系统而发愁。

“非常幸运,阿里云赶上了中国数字产业发展周期和数字中国战略,让我们有机会了抓住移动互联网的机遇、千行百业数字化的机会。”张勇直言。

# 大模型竞速绝非“大厂游戏”

4月11日,阿里云大模型通义千问正式官宣。近日,大模型领域“新兵”不断涌现,王小川的百川智能、商汤科技“日日新”、昆仑万维“天工”……无论是否与互联网领域直接相关,企业们入局ChatGPT的速度与热情丝毫不受影响。越来越多的互联网大佬下场创业,体量较小的企业也纷纷推出自己的大模型。围观者惊叹,原来高投入的大模型从来都不只是“大厂之间的游戏”,大家争分夺秒的到底在争抢什么?



## “诸神之战”

一天之内,大模型领域迎来了三位重量级的嘉宾。

先是搜狗公司创始人王小川,一封公开信正式宣告了“百川智能”的诞生。据了解,百川智能旨在打造中国版的OpenAI基础大模型及颠覆性的上层应用。

紧接着,两家公司先后官宣。在当天的技术交流日上,商汤科技董事长兼CEO徐立展示了“日日新”体系下的语言大模型“商量”,据悉,这是一个1800亿参数的中文大语言模型,可实现多轮对话、逻辑推理、语言纠错、内容创作、情感分析等。

与此同时,昆仑万维也表示,即将推出对标ChatGPT的双千亿级大语言模型“天工”3.5,并定于4月17日启动邀请测试。当晚,昆仑万维发布2022年年报,明确提到“未来10年

将坚定的‘All in’AGI与AIGC。”

大模型成了这一轮AIGC浪潮的兵家必争之地。百度“文心一言”、阿里“通义千问”、腾讯“混元”、京东“ChatJD”、华为“盘古”,传统大厂无一掉队。

垂直领域诸多企业纷纷入局,除商汤科技及昆仑万维外,科大讯飞也在互动平台表示,将于5月6日发布讯飞预训练大模型的详细信息,大模型将用于教育、翻译等领域。

互联网大佬们更不惧大模型的“从0到1”。公开报道显示,王慧文的“光年之外”A轮融资后估值已达到10亿美元,前商汤科技副总裁闫俊杰新公司MiniMax的首款大模型产品Glow累计用户已突破百万。

华西证券研报总结了现有国产大模型的两个特征:一是头部厂商主导,to B为主要模式,赋能原有优势业务;二是均未对公众大面积开放,未有明确的to C入口,实际水平不透明。

## 通用还是垂直

AIGC产业主要分为基础和应用两个层面,前者包含理论框架的搭建、大模型的训练以及算法等,大投入、智力密集和长期性三个特点缺一不可。而后者主要是将别人开发出的东西与自己的业务进行整合,见效快,成本相对较低。

而这也是当下中国AIGC企业押注的两个方向。有人势要成为中国的OpenAI,但也有人将目光瞄准了垂直领域的应用。

天使投资、资深人工智能专家郭涛表示,根据AI大模型应用范围的不同,可以分为通用型大模型和产业型大模型,通用型AI大模型在企业级与消费级市场都拥有广泛的应用场景,如OpenAI旗下的ChatGPT;而产业型的AI大模型主要应用于垂直行业或细分应用场景,如京东旗下的ChatJD。

“整体来看,通用型大模型应用场景相比产业型大模型更为广泛,随着规模化、多样化应用场景的落地,促使算法模型快速迭代升级,不断丰富和拓展新的应用场景,市场竞争也将不会增强。”郭涛认为。

但全联并购工会信用管理委员会专家安光勇则提到,通用大模型在应对多样化任务时具有竞争力,而专业大模型在特定领域有深厚的技术积累和应用优势。

安光勇认为,体量相对较小的企业在在大模型领域与“大厂”抢的主要是技术创新和应用场景的拓展。相较于大厂,“小厂”在技术创新上通常更加灵活和敏捷,能够快速响应市场需求,快速推出具有差异化和竞争力的大模型产品。

此外,小厂在应用场景上可能更加专业化和垂直化,能够深入某一领域或行业,提供定制化的解决方案,满足特定客户的需求。此种专业化和灵活性为小厂在大模型领域与大厂竞争中带来了一定的优势。

“‘大厂更适合做模型’其实是一种惯性思维。”元语智能联合创始人兼COO朱雷解释称,通用的大模型解决不了具体行业的问题,尤其是在国内,大部分高质量中文语料都在行业内,而不在公域的互联网内,即便是大厂也很难拿到行业数据,从这个角度看,专用的模型在特定领域的效果一定要比大公司的通用模型更合适。

## AIGC该降温了

不久前,国际权威咨询机构IDC发布的《2022中国大模型发展白皮书》中预测,2026年

中国人工智能软件及应用市场规模将达到211亿美元,人工智能进入大规模落地应用关键期。

数据显示,自2020年起,中国的大模型数量激增,仅2020-2021年,中国大模型数量就从2个增至21个,和美国量级同等。

萨摩耶云科技集团首席经济学家郑磊认为,在通用大模型方面,中美存在时间差,而应用层则各有千秋。大模型训练成本高,初创企业目前投入的资金很快就会消耗光,“建议政府将其作为新建项目,引导社会资金共同投资”。

此外,郑磊提到,由于还存在可能替代大模型的其他技术路线,国家也应关注,及时加强研究支持,缩小与欧美技术差距。对于中小企业来说,可以根据市场定位和行业发展趋势,选择已有的合适的大模型作为垂直应用研发的基座,逐步形成AI赋能的产品和服务,借助市场力量推动产品研发,形成在细分领域的竞争优势。

上个月末,一览科技推出业界首个基于GPT大模型的“文本+图片+虚拟人”视频AIGC全域工作流,彼时,一览科技创始人兼CEO罗江春就曾提到,AIGC工具只会越来越多,同质化非常严重,说是新一轮“百团大战”也不为过,关键的问题是谁能解决某一个垂直领域用户的根本需求。

“大模型需要降温了。”于资本市场,深度科技研究院院长张孝荣做出了这样的判断。在他看来,大模型是人为炒作起来的一个现象,前期是国际巨头推广产品,后期是国内资本市场跟进,炒作相关企业股价,许多公司股价脱离理性轨道,在短短几个月的时间内翻了数倍,导致资本市场发烧。

北京商报记者 杨月涵