

竞争加剧+成本下降 Token迎降价潮



◎记者 罗茂林

今年以来一路走高的Token(词元)价格迎来反转。进入6月,全球多家大模型厂商开始重新调整Token价格,包括OpenAI在内,多家大厂先后被曝将下调Token价格,以回应企业用户对AI成本暴涨的不满。然而,相较于此前声势浩大的推广潮,此次Token价格下调更像是行业新一轮残酷竞赛的序幕。多位市场人士分析称,随着OpenAI及Anthropic等一批头部AI厂商陆续推进IPO,市场对AI公司的估值模型正在调整,能否用更低成本持续获取更多用户,成为每家公司必须面对的考验。

竞争加剧 Token迎降价潮

进入6月,Token价格开始逆势下跌。日前,头部AI公司OpenAI被曝将于6月启动新一轮调价,计划大幅下调Token计费标准。对于此次降价,外界认为将是一场面向对标竞品Anthropic的行动,目的在于争夺企业级客户。因此,不少业内人士预计,Anthropic或将跟进降价。

事实上,不仅是国际AI巨头,本月初国产大模型DeepSeek正式执行V4系列模型调价,将Token价格下调至原标价的1/4。按照官方公告,DeepSeek-V4-Flash的输入缓存命中价格是每百万Token 0.02元;V4-Pro每百万价格为0.025元。

“这成本几乎比生产对应Token所需要的电费还便宜。”一位资深AI创业者告诉上海证券报记者,作为比较,在今年4月前,

Open AI最新模型每百万Token的综合成本是30美元,“前者价格相当于后者的百分之一,是量级的差别了”。

紧随其后,小米宣布MiMo-V2.5系列大模型完成调价,部分版本降幅高达99%,同时取消按上下文窗口分段计价的规则,并且Token Plan用量直接提升到原来的5倍至8倍,这意味着其每百万Token成本基本看齐DeepSeek。

6月16日,字节跳动旗下火山引擎上线Seedance 2.0 mini视频生成模型,其图生视频定价为每千Token 0.023元,视频生视频定价为每千Token 0.014元,这一定价较Seedance 2.0明显降低。此外,近期腾讯云也发布公告,大模型服务平台TokenHub将于6月15日起下调MiniMax-M3模型价格,推理输入、推理输出及缓存命中费用均下调50%。

有意思的是,今年4月,小米MiMo团队负责人罗福莉曾公开反对打价格战。她认为,低价卖Token并开发第三方框架的行为,看似吸引用户,最终会损害用户体验。

不过,也有技术工程师告诉记者,此轮Token价格的下调与罗福莉的观点并不矛盾:“她强调的是之前模型本身性能不足,不顾成本消耗Token的行为,但现在通过模型和上下文压缩等技术,Token的效率有所提升。”

价格大考来临 行业洗牌加速

从某种意义上讲,此轮降价更像是行

业即将面临的一场商业大考。多位业内人士表示,随着越来越多AI公司步入资本市场,用户规模、实际效益等关键数据的分量越来越重。

“今年以来最明显的变化就是,企业端对于AI的真实需求在变多,它构成了真实的商业订单。”北京某AI大模型独角兽公司公共事务负责人对记者表示,尤其是今年以来Cloud Code崛起,大模型“编程能力”被越来越多企业看到,催生了这一轮Token用量的暴增。

以Open AI披露的数据为例,该公司最大的Token消耗者每月使用量约为1000亿Token,而在6年前,这一数据仅为10万。然而,在企业真实需求增长的同时,Token价格高企的问题日益凸显。据报道,不少美国互联网巨头在近几个月纷纷收紧了对Token使用的预算限制。

记者注意到,不少国内头部互联网公司对Token使用的态度也开始转向:从限制用量到对Token使用仅提供部分报销,显然高昂的成本让不少大公司都有些吃不消。

“模型竞赛基本到了头部决胜阶段,资本需要看到增长,这里面很重要的就是企业端用户规模。”上海某券商计算机首席分析师告诉记者。另一位某头部互联网公司AI负责人对记者坦言,竞争走到这一步,性能之争基本结束,性价比将成为重点。

但有意思的是,对不少头部公司而言,这场降价并非突如其来,相反更像一场预言的兑现。多位技术人士告诉记者,当前影

响Token价格的因素无外乎两个:模型本身技术能力、硬件侧成本控制。眼下,这两个方面都在迅速突破。

在技术侧,对稀疏注意力等技术的持续优化,正在让Token消耗不断降低。以DeepSeek为例,此次最新模型公布的技术突破中就包括压缩稀疏注意力机制(CSA)和高压压缩注意力机制(HCA),这两者都能大幅优化Token在单次任务中的使用。

记者注意到,包括月之暗面Kimi、小米MiMo在内,近期多款国内大模型披露的产品创新性论文中均有类似技术。

在硬件侧,按照业内共识,Token在硬件端最终的变数取决于两方面:芯片成本和电力成本。随着国家“东数西算”工程的持续推进,在不少绿电能源丰富的节点城市,电价和机房租用的成本都在大幅下降。以内蒙古乌兰察布为例,记者调研发现,包括华为、阿里、快手等国内厂商均在加快布局。

某算力中心当地运营负责人透露,由于厂商与算力中心的服务合约中电费占据大头,且多为长年合同,随着电价以及综合服务成本的快速下降,厂商Token生产的实际成本明显回落。此外,多位算力中心人士对记者表示,今年以来国产芯片产能逐步提升,其价格也在回落,带动Token成本进一步下降。

但显然,并非所有的大模型公司都能撑住此轮的价格冲击。有AI行业资深观察人士直言,部分企业实际上是在被动降价。

故事赢人心 科技增效能 ——中国电影寻找“破局”之路

◎记者 杨翔菲 王彭

“观众从来没有远离电影,只是我们从业者还需要做得更好。”

当潮汕方言、全素人出演的小成本影片《给阿嬷的情书》斩获亮眼票房与口碑,当《哪吒》系列凭借200多亿元总票房、数亿人次的观影体量,印证国产动画电影的国民根基,电影行业都在深度复盘一个核心问题:当下观众真正认可、愿意买单的电影,究竟是什么模样?

在第28届上海国际电影节(下称“上影节”)期间,这一问题成为热议焦点。而破局的答案,从来不在被反复推演的大数据模型中,也不在追逐流量、跟风热点的速成创作方法论里。

创作方法论逐步失效

过去10年,国内电影行业试图搭建标准化的创作与市场坐标系,从题材类型、情绪导向到票房预判,各类量化规则、复盘方法论层出不穷。但《给阿嬷的情书》的市场突围,彻底打破了这套固化的行业逻辑,让传统市场预期体系全面失灵。

“过去两三个月,我被问到最多的问题是,《给阿嬷的情书》是怎么被发现的?你们为什么会投?你们什么时候投的?”作为该片的出品方之一,大麦娱乐总裁李捷在上影节金爵电影论坛主论坛上的发言吸引了全场的关注。

据介绍,早在2021年,大麦娱乐便与《给阿嬷的情书》导演蓝鸿春达成合作。2022年贺岁档,双方协力推出其“潮汕三部曲”中的第二部《带你去见我妈》。影片以潮汕方言、家常美食和笑中带泪的亲情感叙事打动观众。这次合作,让大麦娱乐看到了一位青年导演不可多得的作者意识和成长潜力,也埋下了互相信任的种子。

李捷反思道:“过去10年,行业一直在找电影投资的方法论,但大多数方法论都是事后总结出来的。观众永远比创作者审美变化更快,我们曾相信题材和类型能复制,相信大制作、大演员能带来观众。恰恰是每一次好的电影出来,都证明这个方法论可能是错的。”

儒意电影董事长、总裁陈祉希从产业端论证了这一观点。她观察到,在市场紧缩的当下,一些年轻创作者反而抛弃了前辈们“推翻自己,重建剧本”的匠心,转而追逐大演员与大成本,偏离了内容创作的核心本质。

“大麦娱乐始终保持着对精品内容的坚守、对青年导演的耐心陪伴与扶持。正因如此,才能有幸与蓝鸿春导演这样的优秀创作者达成双方认可、彼此选择。”李捷说,“如果我们当年错过了《带你去见我妈》,5年后,就不会有《给阿嬷的情书》。”

以好故事俘获观众本心

当标准化的行业方法论不再适用,什么才是电影行业的“突围密码”?

《我不是药神》的导演文牧野给出了一个极具感性的答案:“电影是关于人的,关于情感的。大到宇宙里,小到胡同里,最后传递的都是人的东西。如果它能传递到我心里,打动我,那就可以。拍电影无非是找到观众,它得有表达欲,有分享欲。它是一个围在篝火边讲故事的过程。”

这种“篝火边的真诚”,恰恰是当下电影市场最稀缺的。李捷表示:“观众现在的阈值大幅提升,很难被一个短视频或一张海报打动,也很难再为故事以外的噱头买单。观众唯一认的,是这部电影的诚意——是否尊重了我,对得起我花的电影票钱。”

陈祉希用了一个生动的比喻:做导演与做厨师一样。有的餐厅装潢很贵,但大家不爱

吃;有的厨师做出来的菜,符合80%以上受众对味道的认定。导演对于情绪价值的理解、对于爱的理解,如果符合80%观众的观念认可,大概率就会打动观众。

《年会不能停!》的导演董润年则表示,观众主体性的增强,意味着创作者必须真正“把自己投入到生活里面”,不能被算法和朋友圈所包围。

创作者坚守内容真诚的同时,产业端的陪伴与支撑同样至关重要。

“我们深知,每一位创作者,尤其是新生代,不缺想象力,但容易感觉孤独。作为电影创作机构,我们怀着非常开放的心态发掘年轻人,希望他们能够自由表达,同时陪伴青年创作者一路成长。”上影集团董事长王隽表示,扶持新生代青年力量需要耐心和定力。

“作为出品方,我们不会告诉青年导演该拍什么,而是努力守护他们想拍什么。”李捷认为,电影的创作周期需要一年甚至更长,单纯追求流行爆点的创作往往滞后于观众需求。“既然如此,不如回归初心,真诚创作。这是对《给阿嬷的情书》再次印证的市场逻辑。”

技术革新助推长效发展

电影行业自诞生之初,便是艺术与科技深度融合的产物。随着科学技术手段不断创新,人工智能正从实验验证环节进入电影生产流通环节。当代电影人面临的核心命题,是如何让生生不息、历经120余年深厚积淀的传统电影工业,与蓬勃发展的AI技术达成深度的双向奔赴。

厦门大学电影学院院长、北京电影家协会主席、导演黄建新表示,技术不再只是工具,而是讲述未来的新语言。因此,电影人将越来越有能力去讲更宏大的故事、去完成更高工业水准的创作。

紫荆影业(深圳)有限公司副总经理李挺伟认为,AI可贯穿电影剧本至后期全流程,核心作用是提升创作效率、提供创作参考,但核心创意、艺术表达与审美判断仍由人主导。“AI能够助力行业降本增效,改变线性创作模式。AI带来的技术平权,迎来了电影创作者时代。”他说。

“AI将为内容行业带来生产力的革新、生产关系的革新。”智象未来联合创始人、首席运营官王科在接受上海证券报记者采访时说:内容行业将从过去的“人+片场”模式,转化为“人+Agent+算力”模式;AI赋能工业流程,全面推动生产效率,让“更快”与“更省”成为可能;同时打破传统叙事边界,全新内容形态涌现。

AI正在怎样具体赋能影视行业?记者走进位于上海影城的AI片场,四组“影视创作者+AI超级创作者”融合创制团队正在探索未来的全新可能。在这个不到800平方米的空间里,融合了算力底座、硬件设施、产业配套、创作力量、互动体验和超级个体展示等内容,搭建起微型AI创制生态场景。

“影视行业真正渴求的是能够理解专业视听语言的‘垂直行业大模型’。”博纳影业集团影视制作副总经理曲吉小江认为,开发并跑通电影垂直大模型,不仅需要底层视频技术的迭代升级、大规模算力的坚实支撑,也需要将专业影视人才的行业经验与美学积淀,转化为模型生成算法的底层逻辑。

在王科看来,AI令生产变得容易,但精品反而会更加稀缺。它并不会稀释价值,而是进一步放大“好故事、好审美、好创意”的价值。“AI不会替我们决定什么是好电影。它可以生成画面、提高效率、打开想象,但最后要不要这一帧、人物有没有情感,还是要靠创作者。”黄建新说。

从规模扩张到精准布局 多家公司优化锂电投资

◎记者 荆淮桥

今年以来,国内锂电产业链投资开启新一轮调整。上海证券报记者梳理发现,年内已有中科电气、天铁科技、德方纳米、恩捷股份、永太科技等多家A股公司公告终止重大锂电投资项目。与此同时,德方纳米、永太科技等公司又上马新项目,以推进产能结构向高端化发展或者完善区域产业布局。

业内人士表示,受市场竞争格局变化、资金压力增大及预期收益难以达成等因素影响,多家公司选择主动终止或暂缓前期规划的产能扩充项目,以控制投资风险、优化资源配置。

终止项目覆盖四大主材

中科电气6月13日公告称,拟终止2022年规划且暂缓建设的甘眉工业园年产13万吨锂电池负极材料一体化项目、兰州新区年产10万吨锂电池负极材料一体化项目以及2024年规划且暂缓建设的摩洛哥年产10万吨锂离子电池负极材料一体化基地项目。三个项目原计划总投资约为103亿元。

“上述投资项目推进过程中,项目实施条件发生客观变化,公司基于降低投资风险考虑暂缓相关项目实施。”中科电气相关人士对记者表示,结合当前产业规模化、集约化发展趋势,从前规划的单点10万吨、13万吨项目已不再具备规模优势,继续推进该批项目已不是优质产能扩充的最优方案。

天铁科技4月28日发布公告,终止在庐江高新区投资年产10万吨改性石墨负极材料项目,该项目原计划总投资约为14.50亿元。公司称,随着项目推进,并结合当前行业环境与公司经营情况,继续推进项目将显著增加公司资金压力,且预计项目实施的预期收益难以达成。

正极材料领域,德方纳米5月21日发布公告,终止曲靖经开区年产33万吨新型磷酸盐系正极材料生产基地、会泽县年产11万吨新型磷酸盐系正极材料生产基地。两项目合计规划年产能44万吨,总投资约为100亿元。公司称,在推进过程中,项目实施条件发生客观变化,项目至今仍未取得实质性进展。为优化资源配置,降低投资风险,决定终止上述投资项目。

在电解液、隔离膜等环节,一些公司也在主动收缩产能。3月3日,永太科技发布公告,终止邵武永太高新材料有限公司年产20万吨电解液及相关材料和副产物循环利用项目。该项目于2022年8月首次披露,总投资额为9.5亿元。公司称,该项目自筹备以来,市场竞争激烈,行业竞争格局发生了显著变化,为提升公司资源配置效率,完善公司生产基地整体布局,公司决定终止该项目。

恩捷股份5月22日公告称,公司董事会同意终止在马来西亚新设子公司投资建设锂电池隔离膜项目。该项目原规划建设产能约为10亿平方米/年,总投资额预计约为20亿元。截至目前,该项目尚未建设投资。

从“跑马圈地”到“精准布局”

从上述公司的公告表述来看,市场环境变化、实施条件变化、优化资源配置等都是高频词汇,反映出锂电行业正从规模扩张向精准布局的变化。同时,相关公司终止的项目几乎都处于前期筹备阶段,尚未实质性建设。在大规模资金投入前,相关公司及时止损,反映出行业投资决策的理性回归。

此前,中科电气在回复深交所年报问询函时表示,负极材料吸引大量企业裹挟资本涌入新能源赛道,存在一定的过度投资和低效产能重复建设等问题。负极材料行业面临产能结构性过剩和竞争加剧的风险,但打通石墨化、碳化及原材料各环节,拥有负极材料一体化产能布局的企业将获得成本优势和竞争优势。

在集中清理久悬的暂缓项目的同时,中科电气表示,公司将集中资源,重点推进泸州年产30万吨项目与阿曼海外基地项目。公司经营布局和发展方向更加清晰,体现出头部企业聚焦核心优势区域的战略选择。

在终止福建邵武项目的同日,永太科技宣布拟与江苏滨海经济开发区沿海工业园管理委员会签订《项目投资协议书》,投资建设年产20万吨锂电池电解液项目,项目总投资约为5亿元。永太科技认为,项目建成后,将进一步完善公司在该区域的产业布局、深化地理空间协同效应,从而实现对该地区及周边市场需求的快速匹配和响应,强化北方区域辐射能力,构建南北联动、资源互补的供应体系,提升整体运营效能,巩固并扩大公司在锂电池材料领域的产业规模。

类似的还有德方纳米。在摒弃老旧产能的同时,公司推进产能结构向高端化发展。公司5月21日公告,拟在禄劝彝族苗族自治县投资新建绿色磷基新能源材料产业链一体化项目,总投资为63亿元。该项目包括年产30万吨磷酸盐正极材料项目和年产27万吨磷酸盐项目。公司称,本次对外投资有利于抓住新能源汽车动力电池和储能市场的发展机遇,进一步扩大产能规模,打造一体化成本优势,满足下游客户需求。

翔丰华4月22日发布公告,拟定向增发募资6亿元用于年产9.3万吨负极材料项目,项目总投资为13.6亿元。公司表示,当前产能和在建产能已有明确的销售意向,为了匹配下游客户的扩产节奏,公司需要适时扩充产能,同步完成产线认证、产品认证,以及绑定头部客户、继续锁定长期订单,持续增强客户黏性,强化行业竞争力。