

城市金融报

城市金融网

www.csjrw.cn



中国馆亮相2024年法国“科技万岁”展

5月22日,2024年法国“科技万岁”展在巴黎登场。今年的展会为期四天,5月22日至24日面向媒体和专业人士开放,5月25日面向公众开放。这是欧洲最大的科技展会之一。图为中国馆内的位形空间科技有限公司提供的虚拟沉浸式互动体验。

新华社供图

AI大模型燃起价格战

金头条

工行台州分行协办

本报综合报道 国内大模型突然进入价格战模式。

近期,国内多家大模型厂商纷纷宣布降价策略,推动大模型从“以分计价”的定价模式向更“以厘计价”新时代迈进。5月21日,阿里云宣布通义千问大模型降价,其中GPT-4级主力模型Qwen-Long降价幅度甚至达到97%;另一大模型头部企业百度,宣布文心大模型的两大主力模型ERNIE Speed和ERNIE Lite全面免费,立即生效;日前,字节跳动在火山引擎原动力大会上发布豆包大模型,并喊出“0.8厘就能处理1500多个汉字,比行业便宜99.3%”的口号。

那么,大模型价格为何“卷”起来了?价格战燃起,谁得益、谁受伤?

大模型价格战愈演愈烈

在持续了一年多的国产大模型鏖战后,云计算大厂终于陷入以价格换市场的老套路中。最近这段时间,大模型厂商们是分不刷了,参数不比了,甚至连长文本也懒得卷了。一个个的车都掉转了车头,打起了价格战。“连降99%”,“注册就送”,“限时免费”等等各种618专属名词,都挂在了各个大模型厂商的主页。

可追溯的源头似乎是5月6日幻方量化旗下DeepSeek发布的DeepSeek-V2,其价格做到GPT-4-Turbo的近百分之一。5月11日,智谱接棒,个人版GLM-3-Turbo调用价格下降5倍。

掀起高潮的主力则是云厂商。5月15日,字节豆包主力模型将其在企业市场的定价做到0.0008元/千tokens,喊出较行业平均价格便宜99.3%,而在5月21日,国内两大头部厂商阿里和百度也入场了。

阿里云宣布:“通义千问9款模型齐降价。GPT-4级别主力模型Qwen-Long价格直降97%,击穿全球底价。”百度智能云则直接宣布文心大模型两大主力模型ERNIE Speed和ERNIE Lite免费开放。

这一波大模型降价潮已经开始出现卷入更多厂商的迹象。5月22日,科大讯飞宣布讯飞星火Lite API(接口)永久免费开放。讯飞星火顶配版API价格不足百度文心一言ERNIE-4.0和阿里通义千问Qwen-Max的五分之一。

5月22日,腾讯云也宣布多款混元模型免费或降价,其中混元-lite模型免费,混元-standard-256k输入价格下调87.5%至0.015元/千tokens,输出价格下降50%至0.06元/千tokens。万亿参数模型混元-pro,API输入价格则从0.1元/千tokens降至0.03元/千tokens。

同样的“价格战”,也在国外开打。OpenAI在5月13日发布的GPT-4o,不仅在功能上大幅超越GPT-4 Turbo,价格只有其一半,为5美元/百万tokens,而2023年年初以来,OpenAI也已进行了4次降价,而谷歌旗下Gemini大模型系列中,Gemini 1.5 Flash的价格为0.35美元/百万tokens,比GPT-4o还要便宜得多。至此,国内外科技巨头围绕AI大模型的价格竞争已经全面展开。

那么,AI大模型为何要降价?

按照阿里云的说法,此次降价是希望加速AI应用的爆发。调价后,未来大模型API的调用量会有成千上万倍的增长。

但据了解,对于降价的目的和作用,业内有不同看法。新浪微博新技术研发负责人张俊林认为,价格低才能促进应用发展,因为便宜所以会有更多应用。

一名国内互联网大厂大模型开发人员表示,大模型厂商此时降价,是希望用低价鼓励用户在更多场景探索落地。现阶段大模型厂商而言,通过定价赚钱不是最主要的因素,最重要的是先把应用场景铺开,把流量做上去。

谁得益、谁受伤

大模型降价后,这些厂商能否抢夺更多用户、培育起蓬勃的应用生态还有待观察。仅从算力资源来看,字节、阿里、百度作为云厂商不需向外购买算力,讯飞、字节、阿里等厂商资金实力也较创业企业雄厚,降价显得更有底气。

不过,随着降价潮继续,做基础大模型的创业企业或许不得不面对趋于激烈的竞争环境。张俊林表示,现在已到了通过价格战淘汰掉一些大模型选手的时候,预计今年下半年还会继续淘汰一批。如果是创业做应用,大模型降价是件好事情,但如果做的是基础模型,则会很难。如果做不出特色或者应对竞争被迫跟着降低成本,会很困难。

在互联网大厂和创投机构有多年从业经历的王先(化名)表示,大模型创业公司拿着技术找应用,要找对方向并不容易,就像做AI代码助

手这类产品,现在可能是可行的,但当互联网大厂把产品延伸到这个领域后,很快就能打败创业公司。创业公司不能跟大厂直接对抗。大模型创业公司要形成壁垒和完整的商业链条,把现金流做起来,这时候才能说是安全的,否则在市场竞争中永远不安全。

猎豹移动董事长兼CEO傅盛则认为,这次大模型降价基本宣告了大模型创业公司必须寻找新的商业模式。降得最凶的都是有云服务的大公司,(他们)通过大模型来获取云用户,羊毛出在羊身上,降得起,而大模型创业公司没有这样的生态,必须另寻商业模式。

业内专家:AI应用有望加速普及

在业内看来,本轮大模型降价潮是各家厂商为了抢占市场份额,加速AI应用的商业化落地而采取的策略。

谈及用户场景落地情况时,火山引擎总裁谭伟举例道,随着汽车与手机的交互日益自然,如在驾驶中导航、查询周边信息,无需分心操作手机,车载系统可即时响应需求。手机助手AI的应用受限于技术,未充分满足用户需求。在其他场景如消费营销、金融服务中,智能体也展现出巨大潜力。为实现AI技术的广泛应用,需提升技术水平并构建丰富的应用生态。这需要低代码生态的支持,降低开发门槛,激发群体智慧。同时,降低AI技术成本,使更多初创企业和小规模项目得以实施,促进AI时代的创新与发展。

北京一位人工智能行业人士表示,本轮大模型价格战的开启,对于AI应用的普及具有积极的推动作用,将使得更多企业和个人开发者能够负担得起AI技术的使用成本,从而更加积极地投入到AI技术的研发和应用中来。价格战的竞争将促进AI技术的不断创新和优化,从而推动整个行业的技术水平不断提升。随着AI应用的普及,将带动相关产业链的发展,形成更加完善的AI生态系统,为整个社会的数字化转型提供有力支撑。

业界普遍认为,随着大模型性能逐渐提升,AI应用创新正进入密集探索期,但推理成本过高依然是制约大模型规模化应用的关键因素。

几次工业革命的成就已告诉人类,科技变革本质上都是伴随着边际成本的大幅下降,进而推动了技术的扩散。有业内人士表示,AI大模型纷纷降价一方面有助于大众普及AI,另一方面也是顺应市场需求。

今日评

工行嘉兴分行协办

理性看待“网约车市场饱和”

□王庆峰

近日,“网约车市场饱和”的话题再度引发关注。江西景德镇、苏州、重庆等地的交通运输局陆续发布了关于网约车行业的风险预警,福建莆田、河南商丘等地也公开提醒慎入网约车行业。

网约车市场到底怎么了?数据最有说服力。根据多地公布的情况,在众多网约车平台中,有的平台没有开展经营,有的平台接入车辆低于50辆,有的平台日均接单量小于1000单;网约车从业人员的经营状况同样不乐观,多地网约车接单日均接单量不超过20单,单车日均营收在20元左右,单车日均收入仅200多元。

其实,这不是最近才有的情况。从去年开始,网约车市场就步入了“供给过剩期”,其中最主要的原因就是网约车及从业人员数量快速增

加。根据交通运输部的数据,2022年初,网约车驾驶员证的发放数量为398.8万本。到了2023年8月,这一数字已经上涨到605.9万本。截至今年3月底,各地共发放网约车驾驶员证679.1万本。也就是说,短短两年多,网约车司机数量增加了约70%。如此“人海汹涌”,市场必然供大于求,僧多粥少。

各地发布风险预警,提醒企业和个人认真调研市场,深入了解政策,谨慎投资市区网约车市场,应该说初衷是好的,可有效保护现有从事网约车行业的司机和企业的利益。但也必须承认,网约车因工作时间灵活、门槛不高,对很多人来说是过渡性选择,是不得已的“人生退路”。国家发改委曾发文提到,外卖骑手、网约车司机等都是吸纳就业的“蓄水池”。如今,这些“蓄水池”也快要满了,相关部门更应该做的是挖潜、扩容,而不能只是开源、节流。

同时也要看到,网约车是一个充分竞争的市场,应该由市场自行调节。根据媒体报道,一些地方不只是单纯发布饱和和预警,而且明里暗

里采取了冻结措施。比如三亚、长沙等地,曾有一段时间暂停受理网约车运输证新增业务,还有很多城市没有明确宣布,但都在实际上限量审批或者干脆不批,以关掉新增市场运力的闸门。这些试图减少市场供给的做法,是非常不理智的,不但损害了更多灵活就业从业人员的权益,也不利于行业健康有序规范发展。

面对“供给过剩”,正确的做法不是减少供给,而是优化供给。比如,现在不少资质不过关的车辆混入网约车队伍,造成网约车市场“观感饱和”的局面,这就应该发挥监管部门作用,加快清退不合规车辆和人员,消除行业风险隐患。此外,一些平台和企业走入了“卷”成本的死胡同,互相比拼“一口价”“特惠价”“高额补贴”等低价营销手段,非但没有提升服务水平,反而恶化了乘客体验,这一情况也需要引起重视,持续抓好服务优化,扩大需求。

综合各方面来看,“网约车市场饱和”是一种善意的提醒,其中蕴含的深层次矛盾,也需要我们加以正视,努力改进。

财发布

网络拍卖领域首个规范性文件出台

近日,市场监管总局、商务部、国家文物局三部门联合印发了《关于促进网络拍卖规范健康发展的指导意见》(以下简称《指导意见》),旨在进一步规范网络拍卖市场秩序,促进拍卖行业高质量发展。

市场监管总局有关负责人表示,《指导意见》是网络拍卖领域第一个规范性文件,填补了网络拍卖管理的制度空白,首次明确坚持线上线下拍卖监管制度的一致性,要求从事经营性网络拍卖应当遵守现行拍卖法律法规。此外,《指导意见》首次明确了网络拍卖平台经营者的主体责任。根据推动传统产业改造升级、支持现代服务业发展的有关要求,《指导意见》首次提出鼓励网络拍卖创新发展的具体意见,其出台有利于促进网络拍卖行业的持续健康发展。

近年来,我国网络拍卖行业蓬勃发展。截至2023年,各类网络拍卖活动的年成交规模已超万亿元,成为非标资产交易的重要方式,在盘活存量资产、提高要素配置效率、完善现代流通体系等方面作用明显,对稳增长、促消费、扩就业有积极贡献。

网络拍卖平台经营者是网络拍卖活动的重要参与者。该负责人表示,平台具有“市场”和“企业”双重属性,在网络拍卖活动中发挥着独特作用,是推动网络拍卖规范健康发展的重要参与者。压实平台主体责任,强化平台内部治理,对保护网络拍卖当事人和平台内网络拍卖经营者的合法权益,营造安全、放心的网络拍卖市场环境意义重大。

因此,《指导意见》对压实网络拍卖平台经营者责任作出一系列规定,主要包括:一是对平台内网络拍卖经营者的身份、拍卖许可、文物拍卖许可等信息进行核验、登记;二是对平台内的违规拍卖信息进行依法处置、报告;三是针对文物等特许经营的拍卖品类,网络拍卖平台经营者要完善专业审核能力和内控制度建设,加强对违法违规交易的动态巡查和及时处置;四是平台内网络拍卖经营者依法履行信息公示义务提供条件,督促网络拍卖经营者公示经营许可相关信息;五是网络拍卖平台经营者作为拍卖人开展的网络拍卖业务应当标记“拍卖自营”,以确保拍卖相关当事人能够清晰辨认。

规范监管和促进发展并重是《指导意见》的突出特点。《指导意见》明确,鼓励各类经营主体参与网络拍卖市场的模式创新。引导各类经营主体参与大数据存储、网络与信息安全维护等网络拍卖基础设施建设。鼓励各类经营主体开发涵盖网络拍卖公告、拍品展示、拍卖师主持、网络竞价、成交确认、电子签署、拍价款结算、资料存档等功能的全流程网络拍卖系统,实现拍卖法定程序功能的全流程线上化。

针对网络拍卖特有的平台主体,《指导意见》鼓励技术创新,引导网络拍卖平台经营者依照拍卖法律法规及业务标准,主动参与网络拍卖市场体系建设,形成良好的网络拍卖市场环境。

针对网络拍卖行业的高质量发展,《指导意见》鼓励服务创新,引导拍卖企业进一步发挥市场和专业优势,创新服务模式,丰富服务场景,拓宽服务内容,促进网络拍卖专业化、规模化发展。针对释放网络文物拍卖发展活力,《指导意见》鼓励业态创新,推动制定网络文物拍卖相关标准,引导网络文物拍卖经营主体延伸产业链条,推进网络文物拍卖与托付保管、在线展示、教育研学等业态融合发展。《指导意见》是首次提出鼓励网络拍卖创新发展的具体意见,其出台有利于促进网络拍卖行业的持续健康发展。

姚君

微博汇

广西北部湾银行协办

人工智能个人电脑产品相继面世



近日,微软全新人工智能个人电脑面世,引起业界关注。近一个月以来,微软、华为、戴尔等企业相继推出人工智能个人电脑新品,抢抓新赛道机遇。中国电子商务专家服务中心副主任郭涛表示,人工智能个人电脑产品目前处于发展初期,各企业都在探索最优的集成方式和最佳的用户体验,这一“摸着石头过河”的过程,将推动新技术与软硬件生态系统的融合。全球头部企业竞相推出新品,无疑将加速这一领域的创新,给产业链上下游带来新的发展机遇。

深圳人才引进入户全流程网上办理

5月22日,深圳人社、公安、发改等四部门联合推出的“人才引进入户一件事高效办”服务正式上线运行,实现人才引进入户全流程网上办理。首批服务适用于应届毕业生、留学回国人员群体,足不出户,最快半小时即可办理入户深圳手续。据了解,外地人才入户深圳涉及人社部门“资格核准”、公安部门“入户指标申领”以及公安部门“户口迁移”等多项子业务。此次推出“人才引进入户一件事高效办”,通过人才引进入户业务办理系统与全国户籍人口管理数据库对接共享,实现数据无缝有效调取。



北京开展数据知识产权试点



5月22日,北京市知识产权局副局长周立权在“两区”建设新闻发布会上表示,北京成为全国唯一被国务院批复开展数据知识产权试点,且是全国唯一全面完成规则制定、登记实践、保护案例、交易使用、数据资产人表等试点工作的地方,试点工作经国家知识产权局验收认定为优秀。目前,北京市数据知识产权登记系统共接收157项数据知识产权登记申请,108项经公示无异议后获得数据知识产权登记证书。其他各项数据知识产权促进运用保护工作顺利推进。