

巨头相继减产后,存储产业二季度筑底

随着上游存储原厂纷纷祭出稳价、减产举措,二季度业绩显示,存储行业陆续开始出现环比向好走势。多家存储头部厂商都表示,行业已经开始筑底。

业绩会上存储厂商普遍提到,其PC和手机终端客户的库存情况已经得到持续改善,有分析认为,伴随下半年消费电子行业旺季到来,存储产品价格也有望企稳回升。

近日,Counterpoint 研究副总监Brady Wang表示,预计存储产业在今年第三季度依然微幅下跌,第四季度将出现持稳表现。“目前跌幅不会出现此前多个季度双位数下滑的表现,但真正稳定还是要等四季度。”他指出,AI行情旺盛也驱动了对存储器的相应需求,但存储器市场最大的需求还是来自手机和PC等市场。因此存储行情的扭转程度,仍要看这类主体市场的需求回升进度。

二季度存储行业环比普遍上涨

近日,全球头部存储厂商相继发布今年二季度财务报告,其中显示环比上一季度,存储行业普遍出现不同程度上涨,当然若要比基数较高的2022年,则依然有较大幅度下滑。

但“筑底”是存储厂商在近期描述的关键词之一。

存储产业Top1厂商三星的二季度财报显示,公司存储业务部分的收入环比微涨1%、同比依然大幅下挫57%。但公司方面表示,产品库存于5月触顶后开始下降,伴随库存调整接近尾声,预计2023下半年需求将逐渐恢复。公司将扩大DDR5、LPDDR5X、HBM等高附加值产品的销售。

存储产业Top2厂商SK海力士二季度财报显示,期内实现营收7.31万亿韩元,环比大幅增长44%,同比依然下滑47%。公司方面表示,内存需求从2023年二季度开始复苏,但不足以使库存恢复正常,2023年DRAM和NAND产能均将下降。由于生成型人工智能开发和商业化的竞争加剧,第二季度对人工智能相关内存产品的需求激增,扩大了高密度DDR5等高端产品的销售。

DRAM行业Top4的南亚科技高管在财报会上表示,2022年二季度开始,DRAM需求逐季下滑,但伴随DRAM供应商在持续调节产能和产品组合配置,2023年二季度行业开始筑底,下半年有机会需求逐步改善,预期第四季供需可望恢复平稳,但仍要观察全球经济景气度的复苏力度。

NAND Flash主控芯片头部厂商群联电子,其执行长潘健成分析,近期有出现部分NAND控制芯片急单需求,甚至发生部分订单有短缺供给的问题,显示市场陆续出现回温的信号。

根据潘健成分析,随着进入传统第三季备货旺季,再加上NAND价格处于相对低档,近期有看到NAND储存模块需求有缓慢增加,包含客户逐渐导入大容量储存模块产品。由此也反映在群联6月Bit Growth,约有36%左右的年增率。

只是从产品线来看,头部厂商的DRAM部分业务明显受益于生成式AI技术对高端存储产品



的旺盛需求;相比之下,NAND Flash则依然要被考虑需要酌情减产。

三星方面表示,下半年将维持下调产量的计划,正选择性地对部分DRAM和NAND产品的生产进行调整,特别是NAND产品,不过公司并未具体说明;同时公司希望提高盈利能力,因此将提高高附加值产品份额。

SK海力士方面有类似思考,公司方面表示与DRAM相比,NAND具有较高的行业库存水平、较低的盈利能力,因此决定进一步削减NAND产量,对其额外减产5%-10%。

当然,预期向好是一方面,最大的不确定性还是来自全球宏观环境变化是否能如预期。

潘健成指出,NAND原厂自第二季开始陆续释出逐渐涨价的讯号,包含7月启动新订单调涨价格或既有订单重新议价,代表NAND原厂不堪亏损的压力持续扩大;然而是否能因此带动市况反转,仍须看整体需求以及市场的接受度。

南亚科技方面则分析,整体供给面,预期2023年产业供给将同比减;需求方面,服务器市场有AI服务器对云端的需求,但企业的IT支出保守,因此美国云服务商的复苏力度需要重点观察;手机在下半年将陆续迎来新机上市,平均DRAM搭载量将增加,但也需观察中国市场的手机构销售复苏力度。

AI成为持续高成长需求的细分市场

在等待需求恢复进程中,AI无疑成为难得持续高成长需求的细分市场。

生成式AI对于存储的需求逻辑在于:算力成长的同时,伴随着对于存力和运力的需求,要将数据进行运输和留存。算力背后主要是GPU支撑;存力则主要是存储器,尤其是DRAM内存的需求;运力则是由于在数据计算和存储之间存在“内存墙”,因此需要如DSP等芯片进行数据运输,这也是近两年来存算一体、存内计算提出的原因,旨在缓解内存墙对算力速率的影响。

由AI服务器带来对存储的需求也在财报会

上被着重关注,HBM(高带宽内存)是其中需求最为旺盛、技术门槛最高的技术。目前三大存储原厂三星、SK海力士、美光均有相应产品配置,以SK海力士和三星占据几乎垄断的90%市场份额。

美光在5月发布公司第三财务季度报告中曾着重提到,AI服务器中对DRAM容量的要求,是通用服务器需求的6-8倍;NAND方面则是3倍。因此生成式AI驱动对AI服务器的需求,成为公司业务的强劲支撑。

SK海力士的财报显示,受益于用于AI服务器的HBM和高密度DDR5模块销量大幅增长,为第二季度的DRAM出货量和ASP增长做出了贡献。

在业绩会上SK海力士高管则表示,包括HBM产品在内的图形DRAM销售额此前仅占公司DRAM销售额的个位数百分比,但在2022年四季度达到DRAM销售额的10%,该比例在快速成长,到2023年二季度为20%以上。公司预计HBM叠加图形类产品,共计占DRAM总收入20%,预估2023年该部分销售额同比增长2倍以上。

但需要注意的是,AI催生的需求并非惠及所有厂商,从大存储板块来说,HBM等高端存储需求依然处在较小的市场规模。

群智咨询副总经理兼首席分析师陈军表示,HBM用于高端GPU和服务器的DRAM,全球一半的HBM供应主要来源于SKhynix,三星次之。由于HBM价格相对较高,服务器市场普遍还是以DDR4和DDR5为主,HBM在整体DRAM中的占比仍然很低。

“AI服务器所需的HBM存储占DRAM体量较低,大约在单个位数比重,因此目前阶段对存储行情的拉动效果不明显;而存储还有接近一半市场是NAND Flash闪存,这部分仍然还有需求下滑的可能性。”Brady也有类似观点,他指出,因此AI服务器的旺盛在今年内还无法为存储器市场带来明显反转,大约在年底才能看到行情一定程度修复。

供稿:《21世纪经济报道》

大模型落地, AI如何颠覆内容生产方式?

人工智能浪潮中,AIGC正在颠覆现有的内容生产方式,以十分之一的成本,以百倍千倍的速度生成AI原创的内容,它不仅会提升内容生产的效率,也会创造出独特的价值和独立视角的内容,从而带来内容生产、内容消费模式的突破。

AIGC技术可重塑数字人

“AIGC技术的发展,尤其是大模型技术的出现,让娱乐行业拥有了大脑和灵魂,而且带来了需求和供给的双增量。”近日,360智脑总裁张向征表示。

以虚拟数字人为例,AIGC技术可以重塑数字人,释放新需求。张向征认为,大模型本身有强大的内容理解能力、用户需求理解能力和文本生成的能力,可以让传统的数字人不再是工具人,而可以和人进行实时对话,满足人的各类需求,解决人的各种工具性问题。

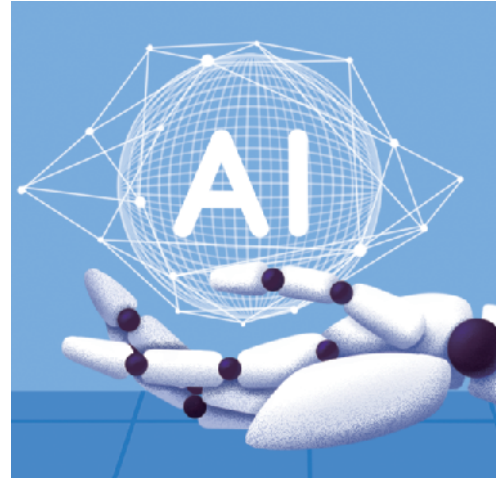
这样的“智能”用到任何一个行业都将是对原有模式的颠覆,自ChatGPT火爆之后,大模型的落地应用就备受关注。

出门问问副总裁李维表示,大模型在创意和生成方面非常擅长,如辅助写作、文案生成等,落地前景清晰可见;但因其本身仍存在的“幻觉”等问题,在对真实性和准确度要求较高的行业中应用仍存在诸多局限。

“在创作类型中,大模型起了很大作用;但在其他场景,大模型的‘幻觉’问题仍需要人去把控。”李维表示。

AIGC技术提升内容创作效率

目前多家公司推出的AIGC应用都聚集在内容创作领域,并且取得了不错的成效,如现象级的AI绘画应用Midjourney,在短短



半年间用户数量飞速增长超1000万,年营收1亿美元。

百度知识图谱部总监余俏俏认为,AIGC技术提升了内容创作的效率,降低创作的成本。更重要的是,可以为用户提供丰富多彩超出想象的创意,并能够助力人们把自己天马行空的想象付诸实行。

“如果需要写实的海报,可以不用找专人,不用布景拍摄就能实现;小孩子天马行空的想象也可以快速变成绘本;非设计背景的人如果工作中需要设计图样,也可以借助AI实现;短视频创作者可以更加聚焦精力在深度洞察和特色打造上,因为AI可以在全流程帮助他。”余俏俏表示。

而对于专业人士,AI同样可以发挥作用。数科技首席执行官唐家渝认为,AI的发展伴随着个人创造力的提升,通过AI工具帮助专业人士激发他们创作出更好的艺术作品的的能力,同时边发的交互体验应当向多边体验发展,实现多人之间的互动。

会“用”大模型变得越来越重要

眼下,大模型凭借其超强的生成能力,已经在快速渗透进各行各业的内容生产中,人类工作者和AI大模型的关系变得微妙。

“对于商业化的内容创作,大模型基本是可以完成的。”李维表示,一些行业受到冲击是不可避免的。

以音乐领域为例,北京酷斯塔夫文化发展有限公司董事长宋柯认为,AIGC对于行业的颠覆是全方位的。“从作词、作曲到编曲,甚至歌手的声音、歌手的形象、虚拟演唱会等等,全盘可复制的话会对原有产业链带来巨大的冲击。”他认为,这次AI带来的冲击可能要比MP3时代给音乐领域带来的冲击更大。

不过,在微软大中华区Azure事业部总经理陶然看来,AI、AIGC、大语言模型的能力作为先进技术是可以辅助人类作为AI“创作副驾”CoPilot的功能,但是AI永远不会替代人类创作者和开发者。

李维也表示,商业性的内容创作者需要快速拥抱大模型,会“用”大模型正在变得越来越重要,这也将渐渐成为一项行业必备技能。

“如何刺激大模型生成更好、更符合要求的內容是一种能力。”李维表示,“这可以在经过短期训练后习得。”

余俏俏认为,未来,大模型与知识技术将推动AIGC效果有质的提升。在算法层面,大模型将进一步统一更多模态、任务、场景;在算力层面,将硬件利用率进一步提高,并且形成更高性能的能力;在数据层面,将从之前更多强调规模到更强调质量,以及与模型打造更好的迭代闭环;在知识层面,更强调精准、更多形态,AI推理和决策中发挥更重要的作用。

供稿:《21世纪经济报道》

ITMT 快报

vivo推出6纳米自研影像芯片V3

近日,vivo推出全新自研影像芯片V3。V3首次采用6纳米制程工艺,能效较上代提升了30%。全新设计的多并发AI感知-ISP架构和第二代FIT互联系统,降低功耗并显著提升了算法效果;同时能够灵活切换算法的部署方式,做到V芯片和SoC的无缝衔接。

在V3芯片的加持下,vivo还推出安卓首发的4K电影人像视频,号称能够实现电影级焦外散景虚化、电影级肤质优化和色彩处理;并且还首次实现了安卓平台4K级的拍后编辑功能,用户可以先拍摄,然后在相册中无损调整虚化和焦点位置,实现如电影大师级别的切焦运镜。

vivo还称,和蔡司持续探讨全新的镜头设计,未来将推出符合全新标准“Vario-Apo-Sonnar”的长焦镜头,这是代表最高品质色差控制的APO设计首次应用于移动影像系统,显著提升变焦性能及质量。长焦镜头还采用了浮动镜组技术,浮动镜片组提供了超高的分辨率和对焦能力,可实现出色的微距拍摄效果;同时还具备奶油般的虚化效果。

vivo还积极拥抱AIGC算法,探索高品质虚拟内容创作在人像拍摄场景的应用。它可以实现生成式人像风格,通过vivo手机就可以拍出虚拟与实景结合更为自然和谐的照片,可以让你在一次旅行中通过照片体验到春夏秋冬四季。借助AIGC算法,还可以实现丰富多样的人像属性编辑,如对光照、年龄、发型、表情等进行编辑。

vivo还推出了vivo质感色彩,官方称浓郁的色彩与细腻的色阶过渡,对特定色彩的独立调校,致敬传统经典。目前X90系列已全部实现了vivo质感色彩的升级,X80系列、vivo X Fold2和X Flip两款旗舰折叠手机预计在9月升级vivo质感色彩。

综合

谷歌发布新模型 让机器人更智能

近日,谷歌发布新的机器人模型Robotics Transformer 2(RT-2),搭载了大语言模型的最新进展,且能帮助训练机器人理解扔垃圾等任务。

RT-2是一种“视觉-语言-行动”模型,可以通过在网络上投喂信息和图像,训练机器人执行响应的动作。这使得机器人变得更加聪明,并赋予它们新的理解和解决问题的能力。

例如,如果希望以前的系统能够扔掉垃圾,必须明确地训练它们能够识别垃圾,以及捡起并扔掉垃圾。RT-2则能够从大量的网络数据中转移知识,它已经知道什么是垃圾,并且可以在没有明确训练的情况下识别垃圾。

谷歌宣称,相比以往第一个版本,新的模型让机器人的性能几乎翻倍,准确率从32%提升至62%。

综合

韩国制成防水且透明 柔性有机发光二极管

由韩国科学技术院电气工程学院和国家纳米制造中心科学家领导的联合研究团队宣布,他们使用MXene纳米技术,成功开发出了一款防水且透明的柔性有机发光二极管(OLED)。新材料即使暴露在水中也能发光和透光,有望应用于汽车、时尚和功能性服装等领域。

透明柔性显示器在包括汽车显示器、生物保健、军事和时尚等多个领域备受瞩目。但众所周知,当发生小变形时,它们很容易断裂。为了解决这个问题,科学家们正在对许多透明的柔性导电材料,如碳纳米管、石墨烯、银纳米线和导电聚合物等开展积极研究。

MXene是一种具有高电导率和透光率的二维材料,具有优异的电化学和光电性能,可通过溶液加工实现大规模生产。尽管拥有这些诱人特性,但其电性能很容易因空气中的湿气或水而劣化,因此其商业化备受挑战。

为解决这一问题,研究团队使用了一种封装策略,可保护MXene材料免受湿气或氧气引起的氧化,进而开发出一种寿命长、抗外部环境因素稳定性高的MXene基OLED。新设计的双层封装薄膜,可阻挡水分并具有柔韧性。其顶部还贴有厚度为几微米的塑料薄膜,使其可在水中洗涤而不会降解。

这款基于MXene的OLED,亮度达到1000坎德拉/平方米或更高,即使在阳光直射的户外也可拥有清晰的显示效果。此外,即使在水下浸泡6小时,该OLED的性能也能保持稳定。

研究人员指出,最新研究将成为MXene应用于电气设备领域的指导方针,可应用于其他需要柔性透明显示器的领域。

据《科技日报》