

数十个国产大模型争艳世界人工智能大会 AI大模型火爆,浙江机遇几何

本报记者 谢丹颖 杨一凡

一个几平方米大的白色隔间,几个拉杆箱撑起简易演播台,别上麦克风,坐下后,对面只有一架相机、一台提词器和一名工作人员,这些就是记者体验“数智人生产工厂”的全部内容。

这里是前不久在上海举行的2023世界人工智能大会的腾讯展区,一个看似还没有普通录音棚专业的地方。记者对着镜头、照着“普通话水平测试必读60篇”录了3分钟口播,24小时后一个复刻版的记者诞生,输入文字或语音即可驱动它口播新闻。看着屏幕中连紧张感都别无二致的AI数智人,记者深切体会到了AI的魅力。

ChatGPT的出现,把AI推到一个新风口。“人工智能,尤其是通用人工智能,已经成为当前社会最热门的一个话题,没有之一。”华为轮值董事长胡厚崑说。

逐鹿大模型,会如何左右AI发展?又会给生产生活带来什么影响?在这个以大模型为核心的AI智能新时代,浙江要怎样抓住AI风口,推进经济转型升级?

群“模”共舞

春江水暖鸭先知。世界人工智能大会开幕前一天,等了近半小时,我们才幸运地遇上一辆空出租车。上车一报目的地,司机便熟练地问道:“看人工智能展去?”他告诉记者,在出租车群里,消息早已传开,“这几天到上海世博展览中心周边去,都是这种单子”。

“张江那块的,无人驾驶嘛。”司机年纪不大,对AI了解不少。聊到兴起,他随即做了个决定,“明天不跑单了,我也去看看新鲜。”

今年的世界人工智能大会,展览面积超5万平方米,创历史新高,但三天展程中每天仍然是人挤人。据官方统计,线下参观人数突破17.7万人。

AI潮流,带动众人的好奇心,更推动行业踏浪前行。

作为国内AI界当之无愧的风向标,本届世界人工智能大会的主题前所未有的清晰,无论是在展区还是论坛,无论专家学者还是行业龙头,大家讨论的话题都离不开“大模型”。连普通参观者,话里话外都带着几句“ChatGPT”。大会主办方甚至设立AI大模型展区,大模型参数在10亿以上的参展厂商,超过了30家。

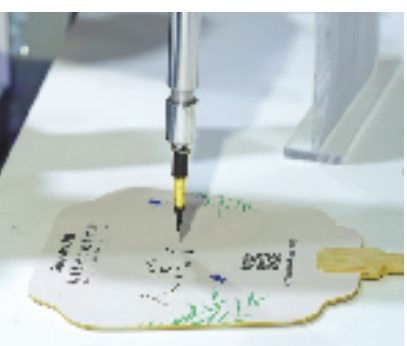
“这是一个令人惊艳的风口。”参会展商纷纷表示,“大模型的发展,至今未见拐点,仍在狂飙式前进。”前脚还在惊叹GPT-3.5的强大,GPT-4就接踵而至,其10秒做出一个网站、60秒做出一个游戏、可准确识别一张“梗图”等能力,再次颠覆大众认知。“要习惯。在人工智能圈子里,半年前的新技术、新风向,现在看起来就算上古产物了,更新速度飞快。”

参会的一名技术人员告诉记者,2018年推出GPT-1,在当时看来简直离经叛道。当时,以面向特定任务的垂类小模型为主流,“讲究造‘专才’”。直到2022年底,ChatGPT,这款基于GPT-3.5的对话式聊天机器人问世,“大家才猛然意识到,‘通才’AI的强大。”

忽如一夜春风来。眼下,百度的“文心”,华为的“盘古”,商汤的“日日新”……各家厂商纷纷谋局落子,可谓“神仙打架”。科技部新一代人工智能发展研究中心5月发布的《中国人工智能大模型地图研究报告》显示,我国10亿参数规模以上的大模型已发布79个,研发的大模型数量目前排名全球第二,仅次于美国。

但火热表象下,国产大模型“好像都是那么回事”,在本届世界人工智能大会,穿过人声鼎沸的大模型展区,不少业内人士产生了这样的看法。

“我们所处的早期阶段,谁都不知道,大模型的边界、上限在哪。”腾讯云智能副总裁、腾讯云智能负责人吴运声指出,无论是发展时间还是技术本身,我国大模型都还很初级,随着模型变大、算力提升,大模型的未来,正如大会腾讯论坛主题——“无限可能”。“只有更多的加入、更多的尝试,才能产生更多的可能性。”他说。



SoundAI 声智机械手臂进行团扇书法写作。



世界人工智能大会论坛现场



大会主办方供图

人形机器人进行“千手观音”表演。

如何落地

一位参观观众这样评价大模型:“相信大多数人和我一样,并不完全理解大模型是什么。”他说,“更像是一种算力硬件支撑下的智能软件。”

但即便是AI“小白”,也能明显觉察一个风向——更多投资人、企业的焦点,开始向垂直大模型或应用层迁移。“未来每一家公司都需要成为数字化公司,每一个应用都将由人工智能驱动,微软将通过AI推动产业升级。”微软大中华区董事长兼首席执行官侯阳在会上强调。

“落地”成为继“发展AI大模型”之后,人工智能圈子的第二个共识。

以腾讯为例,在本次大会上他们高调公布了“行业大模型”的路线,一口气抛出包括金融、文旅、政务等10大行业超过50个解决方案。

记者逛一圈下来发现,展区仿佛一家大模型超市,“大模型+金融”“大模型+教育”“大模型+无人驾驶”……各个行业专属解决方案陈列其中,琳琅满目。

云知声,一家智能语音识别公司,其大模型“山海”,主要聚焦医疗和智能家居,辅助生成电子病历、当好智能管家。

云深处,这家杭州企业带着他们最新款四足机器人亮相大会,实地作揖、蹦跳、翻跟头,吸引不少参会人士驻足互动。记者也去凑了个热闹——双脚踩着“绝影X20”的背,随着它流畅完成提膝、抬脚等动作,并一直被稳稳托起,在心跳加速中感受了一把AI的力量。

问及是否与大模型有结合,“机器人还在学。”云深处科技市场品牌经理钱晓宇告诉记者,下一波AI浪潮,或许便是“专才”AI和“通才”AI的强强联合,“用GPT为机器人注入‘灵魂’”,让它成为能够理解、推理并与物理世界互动的智能系统,实现“具身智能”。

但大模型落地这条路,并非坦途。记者在腾讯展区体验“数智人”时,一位音乐人从北京远道而来,咨询合作事宜。但只聊了几句便发现,腾讯多媒体实验室自研AI作曲框架“Xmusic”并不能直接用在他想要的

未雨绸缪

高昂算力、术语晦涩、人才专业……AI仿佛很远。但在杭州丁桥,AI又很近,融在社会生活的毛细血管里。

曾经,逛超大型超市,对家住丁桥的老唐来说,可不是件容易事:倒一班车,下车再步行300多米,才能坐到大超市的公交。如今,他只要下个楼,搭乘“丁桥小蓝巴”1998路就能直达。

这辆公交是城市大脑利用出行数据等为丁桥居民专门“算”出来的,串联起了丁桥区块的生活区、地铁站、商贸区、学校等绝大部分生活场景。

杭州城市大脑有限公司总经理申永生跟记者算了笔账,如果“丁桥小蓝巴”能为每位丁桥居民平均节省5分钟,丁桥有约40万常住人口,这意味着每年共可以节省200万分钟的时间。

“虽然这个计算有些理想主义,但即便这个数字再缩小到十分之一,也极有意义。”申永生说。

“城市大脑好比未来10年人工智能的‘登月计划’。”中国工程院院士、阿里云创始人王坚曾这样比喻城市大脑与AI两者之间的关系。而大模型的赋能,可谓“登月火箭”的成功“上天”,让“登月计划”的实现又近了一步。申永生对此寄予厚望,他告诉记者,大模型技术BERT赋能社会治理“一网统管”的探索,在识别城市运行中的风险、助力城市高水平治理上,已经取得很好效果。

每一次技术革命,都会推动城市治理的进步,也在加速产业的变革。而之于AI这波大潮,浙江也早有准备。在本届世界人工智



工作人员在2023世界人工智能大会展会现场向参观者介绍百度“文心一格”大模型。

LiveHouse现场。“会后再详谈。”据悉,一天有近10人留下了合作意向并留下了联系方式。正如吴运声所说:“企业需要的,是在实际场景中真正解决某个问题,而不是在100个场景中解决了70%-80%的问题。”想要真正解决客户最核心的业务问题,大模型的私有化、定制化必不可少。而眼下,做一个大模型,部署成本并不低。

但参会展商多数表示乐观:“大模型的核心技术,以及算法、芯片等上游产业链都在快速迭代,成本下降很快。”以百度为例,今年3月首发的大模型,到5月调用文心一言的成本,就已经降低到了刚发布时的10%。

随着新技术的兴起,AI在创投界不断进入春天、冬天,如此循环。每一次看似高蹈的技术理想面前,都悬着一把难以商业化的达摩克利斯之剑。这次也没有例外。



机器人Ginger进行投篮表演。

(本版图片除署名外均据新华社)

能大会开幕的前一天,全省平台经济高质量发展大会上,AI便被划了重点。

“去年以来,人工智能大模型加速走向成熟,牵引各行各业从数字化迈向智能化。”阿里巴巴集团董事局主席、首席执行官张勇指出。

对此,浙江大学人工智能研究所所长吴飞进一步细分了“赛道”,他告诉记者,对于浙江而言,在通用大模型快速发展的同时,一批诸如教育、司法、金融和工业等垂直领域大模型也在打造之中,倒逼这些行业的转型升级,“人工智能不仅发挥头雁效应,也正成为生产力基座。”吴飞说。

联通(浙江)云数据有限公司总经理赵才

苗表达了类似的观点,在他看来,大模型将为浙江产业升级带来巨大的发展机遇。作为民营经济大省,同时又是数字经济先行省,浙江在推动产业链与创新链深度融合上,具有多方面的先发优势。“联通云数据也在数据中心、云网融合、算力赋能等数字化基础设施上,做好了充分准备。”

“最重要的是,传统行业一定要关注AI。”创新技术的发展,未来赋能各个行业后,带来的很有可能是颠覆性的变化。“以后,打败你的往往不是同行,而是掌握了这些最新趋势的其他行业。”2023世界人工智能大会上,大家纷纷表示,未来已来。

记者手记

有热闹亦有烦恼

杨一凡 谢丹颖

ChatGPT的爆发,让AI大模型掀起了一场全球风暴。2023世界人工智能大会三天展会期间,永远人流如潮,我们在第一线,充分感受到了普通公众对于科技进步的热忱与关注。

印象颇深的细节,来自一些大模型展区。我们遇见了一位讲解员领着一个10岁左右的小男孩,尽量用孩子能听懂的话语,边逛边介绍。我们还遇见了一位特殊的看展人,步履轻松、穿着时尚的他,一问年纪已近70岁,问及保养秘诀,他笑道:“多赶潮流、多看热闹,保持对世界的新新鲜感,保持对新事物的学习,自然就年轻。”

我们与很多参展厂商进行了交流,无论是做AI大模型的还是做机器人的,无论是大厂还是初创企业,交谈间,都能感受到他们对于未来发展的自信。

大模型掀起的新一轮AI浪潮已来,但我们也发现,AI狂飙的路上,仍有硬骨头要啃。

问及国内大模型与国外大模型的差距,我们听到了几乎一致的回答:国内在应用层面的发展市场广阔,应用场景和技术层的发挥空间无比广阔,但是在基础层面跟国外相比仍有较大差距。同时,厂商们指出,未来在大模型的落地、赋能上,还有部署成本、算力缺口、模型应用安全问题等“成长的烦恼”。

大模型时代,有热闹亦有烦恼。AI的革命性转变时刻或许还在路上,但不可否认,对中国的AI行业来说,未来已来。

名词解释

AI大模型

即人工智能预训练模型,包含了大算力、强算法、大数据三大要素。AI大模型指一个庞大复杂的神经网络,需要通过存储更多的参数来增加模型的深度和宽度,从而提高模型的表现能力,参数从百亿起步,对大量数据进行训练并产生高质量的预测结果。以众所周知的AI大模型ChatGPT为例,根据创建这一大模型的公司OpenAI的公开数据,早在2020年训练GPT-3模型时,用到的参数数量和训练量便高达1750亿和45TB,消耗算力3640PF-days,即按每秒一千万亿次计算,需运行整整3640天。更不必说参数数量持续扩大的GPT-4模型。

通用大模型和行业大模型

通用大模型是指具有广泛适用性的大型生成式AI模型,能够处理多种任务和领域。而行业大模型则是在通用大模型的基础上,针对特定行业需求进行优化和调整的模型,具备更强的专业性和针对性。

在业务场景中,通用大模型和行业大模型各自发挥着重要作用。通用大模型由于其强大的通用性,可以广泛应用于智能客服、智能推荐、自然语言处理等领域。而行业大模型则能够更好地满足特定行业的特殊需求,比如医疗、金融、教育等。

GPT

是Generative Pre-trained Transformer的缩写,即“生成式预训练转换器”。其可在海量通用数据上进行预先训练,具备强大的语义理解和生成能力。在学习大规模文本数据后,它可以根据用户输入的文本自动生成连贯、流畅的语言,逐渐实现和人类自然语言的无缝对话。这将大大提升人机交互的质量和效率,为各行各业的技术创新带来巨大的潜力,在智能对话、语言翻译等应用领域有广泛的应用。风靡全球的ChatGPT,就是基于GPT再加上人类互动行为之后,所设计的一种AI聊天机器人程序。

