



# ChatGPT 变革与风险

美国人工智能公司OpenAI的大语言模型ChatGPT在推出约两个月后,1月已达到1亿月活跃用户,成为历史上增长最快的消费者应用程序。相关专家预计,ChatGPT不仅是新一代聊天机器人的突破,也将为信息产业带来巨大变革,但由此带来的学术造假、技术滥用、舆论安全等风险亦不容忽视。

## 新一代操作系统平台的雏形

多语言撰写充满想象力的诗歌,编写可运行的程序,快速生成论文摘要,自动制作数据表格,纠正文章中的语法和表达错误,把一周大事写成新闻综述……ChatGPT不仅能理解很多人类问题和指令,流畅展开多轮对话,也在越来越多领域显示出解决多种通用问题的能力。

ChatGPT还轻松通过一些对人类难度较高的专业级测试:它新近通过了谷歌编码L3级(入门级)工程师测试;分别在B和C+的成绩通过了美国宾夕法尼亚大学沃顿商学院MBA的期末考试和明尼

苏达大学四门课程研究生考试;通过了美国执业医师资格考试……业界形容它的诞生是人工智能时代的“iPhone时刻”,意味着人工智能迎来革命性转折点。

“ChatGPT的成功不应仅仅被看作新一代聊天机器人的突破,而应该重视其对人工智能乃至整个信息产业带来的革命。”北京智源人工智能研究院院长黄铁军接受记者专访时说,人工智能领域的过去十年是深度学习的十年,但产业总体上并没有出现移动互联网和云计算级别的爆发,“ChatGPT的出现,具有划时代意义,大模型+ChatGPT已形成新一代操作系统平台的雏形”。

黄铁军说,ChatGPT在技术路径上采

用了“大数据+大算力+强算法=大模型”路线,又在“基础大模型+指令微调”方向探索出新范式,其中基础大模型类似大脑,指令微调是交互训练,两者结合实现逼近人类的语言智能。ChatGPT应用了“基于人类反馈的强化学习”训练方式,用人类偏好作为奖励信号训练模型,促使模型越来越符合人类的认知理解模式。

“这样的AI可帮助人类进行真实创造,尤其是帮助人类提高创造效率,比如提高获取信息的效率或提出新颖想法,再由人解决其真实性问题。创造效率的提高将产生巨大效益和多方面影响,可以改变世界信息化格局。”中国科学技术大学机器人实验室主任陈小平对记者说。

## 引发新一轮人工智能科技竞赛

ChatGPT的问世正在人工智能领域引发新一轮科技竞赛。北京时间2月8日凌晨,微软推出由ChatGPT支持的最新版必应搜索引擎和Edge浏览器,宣布“重塑搜索”。微软旗下Office、Azure云服务等产品都将全线整合ChatGPT。

北京时间2月7日凌晨,谷歌也发布了基于谷歌LaMDA大模型的下一代对话AI系统Bard。同一天,百度官宣正在研发的大模型类项目“文心一言”,计划在3月完成内测,随后对公众开放。阿里巴巴、京东等中国企业也表示正在或计划研发类似产品。

人工智能大模型领域的全球竞争已趋白热化。黄铁军认为,ChatGPT未来有望演变成新一代操作系统平台和生态。这种变革似移动互联网从个人电脑到手机的转化,大部分计算负荷将由大模型

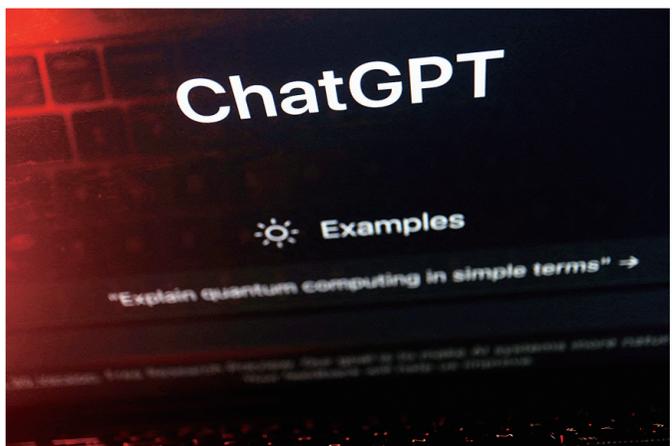
为核心的新一代信息基础设施接管。这一新范式将影响从应用到基础设施各层面,引发整个产业格局的巨变,大模型及其软硬件支撑系统的生态之争将成为未来十年信息产业焦点。

值得注意的是,ChatGPT有时会“一本正经地胡说八道”,存在事实性错误、知识盲区和常识偏差等诸多问题,还面临训练数据来源合规性、数据使用的偏见性、生成虚假信息、版权争议等人工智能通用风险。多家全球知名学术期刊为此更新编辑准则,包括任何大型语言模型工具都不会被接受为研究论文署名作者等。

“学术论文的署名作者须满足至少两个条件,其一是在论文工作中做出‘实质性贡献’,其二是能承担相关的责任。目前这两个条件ChatGPT(以及其他AI系统)都不满足。”陈小平说。

ChatGPT也有应用在舆论信息战方面的潜力。加拿大麦吉尔大学研究团队曾使用ChatGPT前代模型GPT-2阅读加拿大广播公司播发的约5000篇有关新冠病毒疫情的文章,然后要求其生成关于这场危机的“反事实新闻”。连OpenAI也警告使用ChatGPT的用户,它“可能偶尔会生成不正确的信息”,“产生有害指令或有偏见的内容”。

“针对这些问题,我们需要在发展技术的同时,对于ChatGPT应用边界加以管控,建立起对人工智能生成内容的管理法规,对利用人工智能生成和传播不实不良内容进行规避。同时加强治理工具的开发,通过技术手段识别人工智能生成内容。这对于内容检测和作品确权,都是重要前提。”北京瑞莱智慧科技有限公司副总裁唐家渝说。新华社北京2月13日电



这是2月8日拍摄的智能聊天机器人ChatGPT网站页面。

新华社/路透

## 韩国初创企业发布AI芯片

据新华社电 韩国初创企业“反叛”公司2月13日发布人工智能(AI)芯片ATOM。路透社报道,这是韩国拓展AI技术硬件领域的最新动作,“反叛”公司希望这款芯片能中标政府采购项目。

ATOM芯片在机器视觉和聊天应用方面性能突出。公司创始人兼首席执行官朴晟炫(音译)说,鉴于这款芯片主要执行具体任务,而非通用任务,其耗能仅相当于美国英伟达公司A100芯片的约20%。

AI正处全球信息技术行业话题中心。瑞士联合银行集团说,美国开放人工智能研究中心研究的大语言模型ChatGPT,在发布两个月后已成为用户数量增长最快的消费者应用程序。

ChatGPT去年11月发布,因能在数秒内按用户需求写出论文、诗歌或编程代码而引起轰动。业内专家预计,

ChatGPT不仅是聊天机器人的重大突破,也将给信息技术产业带来巨大变革。

AI领域的上述突破同样吸引韩国政府关注。路透社报道,韩国政府希望培育本国AI行业,尤其在AI芯片领域,计划今后5年斥资逾8亿美元投入研发,以期到2030年使韩国AI芯片占本国数据中心市场份额从眼下几乎为零提升至80%。

韩国科学技术信息通信部一名官员告诉路透社记者,本月将就两家数据中心发布招标公告,仅面向本国芯片企业。按照朴晟炫的说法,韩国政府正在强制要求数据中心使用本土AI芯片。

路透社报道的数据显示,全球一半的存储芯片由韩国供应。报道认为,韩国政府希望打造一个市场,作为AI芯片制造商的试验场,以培育具有国际竞争力的本国AI芯片制造商。

## 北京拥有 人工智能核心企业1048家

新华社北京2月13日电 在13日举行的北京人工智能产业创新发展大会上,北京市经济和信息化局正式发布《2022年北京人工智能产业发展白皮书》(以下简称白皮书)。白皮书显示,截至2022年10月,北京拥有人工智能核心企业1048家,占我国人工智能核心企业总量的29%。

根据白皮书,北京核心区域的产业集聚能力较强,已经形成了全栈式的人工智能产业链。北京人工智能领域核心技术人才超4万人,占全国的60%;在专利授权数量全球排名前100的机构中,北京总部机构30家;我国已有24家企业获批建设国家新一代人工智能开放创新平台,其中10家企业总部在北京。

北京人工智能支撑能力

也在不断增强,人工智能算力发展势头强劲,比如北京市政务数据开放走在前列,已无条件开放115个市区级单位,公共服务事项指南、财税金融、城市管理等领域公共数据集15880个;智源“悟道2.0”智能模型的参数规模达到1.75万亿;百度“文心”大模型参数规模达到2600亿。

白皮书表示,今年北京市将引导企业、高校、科研院所、新型研发机构、开源社区等,围绕人工智能关键核心技术协同攻关,持续推动建设具有全球影响力的人工智能创新策源地。同时,支持头部企业打造对标ChatGPT的大模型,构建开源框架和通用大模型的应用生态,形成人工智能产业发展新突破。

讲文明树新风 公益广告 镇江日报社 宣

# 践行社会主义核心价值观

# 人民有信仰 国家有力量 民族有希望

富强 民主 文明 和谐  
自由 平等 公正 法治  
爱国 敬业 诚信 友善

