

当中国遇上 ChatGPT

外行看着很热闹

毫无疑问,ChatGPT一路从国外火爆到了国内。百度、阿里巴巴、腾讯等多家科技巨头纷纷宣布投资开发自己的类ChatGPT人工智能项目,还有一些公司也在有意无意地以“概念股”之名跟风而动,触发接连几周的股价飙升,引发了监管机构对此轮投机性上涨的密切关注。

另一方面,在科技、法律、媒体、版权等多个领域、海内外的专家学者们不断发出警示,让人们必须非常小心应对这头“怪物”,将它称之为科技界的“毒苹果”,教育界“公敌”,本质上是“高科技剽窃系统”“最大假讯息制造源”……

近年来,人工智能技术经历了高速发展,大部分人都已经逐渐接受AI至少在短期内还不会“杀人”或者取代人,但坏消息是人类必须要同AI一起进化,才能在未来新世界生存。

那么,在大家还震撼于ChatGPT的好玩、强大时,中国应该如何应对它的“一本正经的胡说八道”?类ChatGPT技术在中国又将如何演化,成为社会经济发展的助力,以“人机结合”的全新姿态去开创未来呢?

■新快报记者 郑志辉



代注册账号涉嫌非法经营 直接引用涉嫌侵权或违规

去年12月出现的ChatGPT正红遍全球,并成为史上用户增长最快的一款消费类应用。这款聊天机器人可以和用户自然对话,并能回答几乎所有问题,效果十分惊艳。

但必须强调指出的是,所谓惊艳、人性、有趣、狡诈、虚伪等等的评价,都是来自于海外用户和媒体的反馈、报道。通信界人士指出,OpenAI提供的服务应该属于增值电信业务的范畴,如果想进入中国市场,则需要办理增值电信业务许可证,即ICP证。而法律界人士则认为,短期内ChatGPT都不太可能进入中国内地。ChatGPT在对语义的分析中可能会抓取一些涉及到国家安全的信息,同时AIGC(人工智能生成式内容)本身就带有很强的价值传输属性,可能会对价值观和主流观点造成影响,甚至危及国家安全。

这给一些不法分子提供了可乘之机,借ChatGPT热度牟取非法利益。

2月24日,广东省消委会发布消费提示,提醒广大消费者代注册账号涉嫌非法经营,账号可能被收回。广东省消委会相关负责人表示,随意下载相关App程序或注册账号,存在诸多风险。一是些App或应用程序可能会收集和滥用用户个人信息数据,侵害用户个人信息安全权;二是虚假的AI语言模型质量差,响应不准确或不恰当,服务质量低,更可能随时关闭跑路,用户售后服务无保障;三是一些虚假应用程序可能包含恶意软件,进而危及用户设备或数据安全。

除此之外,该负责人还指出,“直接引用ChatGPT答案,可能涉嫌侵权或违规。”因为ChatGPT是由大量数据训练出来的语言模型,在训练的过程中会用到受著作权保护的素材,直接引用可能会涉及侵权等问题。即使不侵害著作权,也可能存在学术伦理问题。目前全球已有多家期刊、出版机构禁止将ChatGPT列为论文合著者,多所高校禁止学生使用ChatGPT。



中国将如何正确应对ChatGPT?

ChatGPT的话题是如此之热,在2月24日国新办的一场新闻发布会上,有记者提问说,近期ChatGPT在国内引起了广泛的反响,科技部如何看待这种人工智能驱动的技术,从监管方面,科技部有何考量?

科技部高新技术司司长陈家昌称,基于自然语言理解的人机对话是人工智能发展的一个重要方向,ChatGPT最近形成了一种现象级的应用,表现出很高的人机交互水平,

表现出自然语言的大模型已经具备了面向通用人工智能的一些特征,在众多行业领域有着广泛的应用潜力。科技部王志刚表示,我们国家对于任何一个新的技术,包括AI技术出来以后,在伦理方面都采取了一些相应的措施,对科学技术发展趋利避害,让其更好地发挥出来。

科技部的回应实际上为我国将如何应对ChatGPT技术带来的冲击作出了定调,关键就在各行各业如

何趋利避害,让新技术造福社会的同时,防范和限制其负面影响。

日前,北京市经济和信息化局发布《2022年北京人工智能产业发展白皮书》中明确提出支持头部企业打造对标ChatGPT的大模型,着力构建开源框架和通用大模型的应用生态。另有消息称,监管机构正出手对ChatGPT掀起的概念股炒作降温,多个AI行业龙头由本月8日起陆续收到“监管工作函”。

复刻ChatGPT有多难?

读者可能要问,打造对标ChatGPT的大模型为何要政府支持头部企业来干?中国这么多大大小小的科技企业、AI企业不行吗?

从技术角度来看,ChatGPT其实质仍然属于自然语言NLP的范畴,是人工智能技术长期发展和积累的结果。Meta首席AI科学家Yann LeCun评论说:“就底层技术而言,ChatGPT并没有什么特别的创新,许多研究实验室正在使用同样的技术,开展同样的工作。”

大模型的训练离不开算法、算力和数据,算法决定了其本身的学习能力,算力则影响着算法的学习效率,数据决定了AI的训练效果。

关键就在于大模型训练的门槛非常高。按国盛证券估算,类ChatGPT大模型的训练成本,一次就要烧掉200万到1200万美元,每日电费高达5万美元;摩根士丹利的分析甚至认为,ChatGPT每生成

一条资讯都可能会花掉OpenAI 2美分,这大约是Google搜索查询平均成本的七倍。还有研究机构估算出,以日均100万用户为基础估计,ChatGPT每日可能要烧掉10万美元、每月300万美元,而自去年12月至今,其访问用户早已过亿。

由于研发所需资金太大,OpenAI在2019年转变为盈利上限型组织,和微软达成了资本挣到有限收益之后,后续收益归非盈利实体。截至目前,微软对OpenAI最新一轮投资额达100亿美元。

这也是为什么日前原美团联合创始人王慧文在社交平台官宣出资5000万美元,要打造中国的OpenAI,却被网友质疑“5000万美元够干什么”的原因。

2月20日,由复旦大学计算机科学技术学院邱锡鹏教授团队研发的中国首个类聊天机器人ChatGPT模型MOSS发布至公开平台,

接受公众内测,不过数小时后就因为访问量过大被网友挤“崩溃”了。

“MOSS还是一个非常不成熟的模型,距离ChatGPT还有很长的路要走。”复旦研发团队21日发布道歉公告表示,复旦MOSS仅为学术实验室,无法制作与ChatGPT能力相近的模型。MOSS团队核心成员之一的张奇教授曾在一次公开讨论中表示,按照ChatGPT目前开源的最简单复现版本、最便宜模式去计算,要做到其1750亿的参数规模,需要大概6000万元的硬件成本,同时运算3.5个月。这还是所有事情都做对的情况下,如中间有参数调整,或者想加速训练过程,就需要更高规模的投资。

到目前为止,国内只有几家科技大厂基本具备训练大模型的能力,比如百度的文心大模型,阿里“通义”大模型,华为的鹏城盘古大模型,腾讯的混元AI大模型等。