



ChatGPT 为何没能诞生在中国?

“

“这是一个‘卡脖子’的问题。”刘知远对《中国新闻周刊》说。目前,ChatGPT 对国内的企业没有开放,相关产业就无法进入到它的体系中。在刘知远看来,OpenAI 已经做了大模型和产品,更重要的是,“我们能不能发明出自己创新的技术和产品”。

聂再清认为,最关键的是信息闭塞,而是国内愿不愿意投入到看上去“无用”的研究中,“现在不少业内人士还是希望直接研发有效有用的东西,不会对一些短期看上去无用的事情上进行大量投入,尤其是像开发 ChatGPT 这样大的投入”。



2月9日,发表在美国《科学公共图书馆·数字健康》的一篇文章提到,ChatGPT 参加了美国执业医师资格考试。这项以高难度著称的考试中,共 350 道题,内容涵盖基础科学、生物化学、诊断推理和生物道德学。ChatGPT 未经专门的培训或强化,经过两名评审员打分,三部分的成绩接近或超过了及格门槛,并表现出了高度的一致性和洞察力。

近两个多月,这款聊天机器人写诗、写小说、敲代码,几乎无所不能,掀起一场 AI 风暴。比尔·盖茨称赞,ChatGPT 将会“改变我们的世界”。但 OpenAI 首席执行官山姆·奥特曼也承认,ChatGPT 虽然很酷,却频繁出现错误信息,经常宕机,是个“糟糕的产品”。

近日的一次采访中,奥特曼说,现在推出的 ChatGPT 只是 70 分版本,还在准备更加强大的模型以及在研究更多东西,“不把这些东西拿出来的原因是,人们会认为我们有一个完整的通用人工智能(AGI),准备好按下按钮,这还差得远”。

有更高“情商”,但不必神化

为与 ChatGPT 较量,当地时间 2 月 6 日,谷歌宣布,将推出自己的 AI 聊天机器人 Bard,并在推特发布了宣传样片。Bard 使用的是谷歌在 2021 年推出的大型语言模型 LaMDA,所用技术与 ChatGPT 相似。

宣传样片中,有人提问,“请问詹姆斯·韦伯太空望远镜有哪些新发现?”Bard 给出 3 个条理分明的回答,但两条是错误的。Bard 称,“韦伯望远镜发现‘绿豌豆’星系的时间是在 2023 年”,这一回答的正解是 2022 年 7 月。

“该望远镜拍摄了太阳系外行星首张照片”的回答也不准确,2004 年,欧洲南方天文台利用智利的超大望远镜已拍摄了太阳系外行星照片。

这款聊天机器人目前仍在内

测,会在未来几周开放给更多公众。首秀“翻车”两天后,谷歌母公司 Alphabet 股价大跌,市值损失 1000 亿美元。

当地时间 2 月 7 日,微软紧随其后发布了新品,推出新版搜索引擎必应和 Edge 浏览器,新增了和 ChatGPT 一样的 AI 语言模型 GPT3.5,微软称之为“普罗米修斯”,并表示比 ChatGPT 更强大。微软演示中,搜索“宜家双人座椅是否适合 2019 年款本田某商务车型”,除传统搜索页面外,页面右侧多了一个聊天窗口,详细列出宜家双人座椅和这款车的车内空间尺寸,给出使用建议,并附上相关链接。该搜索引擎仍在试用阶段,微软 CEO 萨蒂亚纳德称这一天为“搜索领域崭新的一天”,并向谷歌宣战,“比赛从今天开始”。

2022 年 11 月 30 日,OpenAI 推出了 ChatGPT。该公司总裁格雷戈·布洛克曼在接受采访时坦言,“当时并不知道它是否成功”。全球最领先的 AI 语言模型如此直接、低门槛地向普通大众敞开怀抱,并接受所有人的“审视”,这几乎是第一次。

前微软亚洲互联网工程院副院长、现小冰公司 CEO 李笛对《中国新闻周刊》分析说,ChatGPT 之所以出圈,特别之处在于,它被普通用户观察到,并超过了人们对人工智能的预期。

用户尝试用千奇百怪的问题试探 ChatGPT 的边界。它听得懂一段相当复杂的指令,比如,“给 OpenAI 创始人写一封英文采访邮件,探讨对 ChatGPT 使用的理性思考,提到 ChatGPT 的热度及遭受教育界抵触的事实,并列出 5 个问题”。它能指出一段话中的语法和逻辑错误,有人恶趣味地骗它,“我吃了一辆汽车,现在有点饿,该怎么办?”它很快识破,“很抱歉,吃汽车是不可能的,如果您感到饥饿,我建议您吃点实际的食物。”甚至你还可以要求 ChatGPT 在聊天中再构建一个 ChatGPT,看着它和自己聊天。

今年 1 月,美国北密歇根大学哲学教授安东尼·奥曼向

《纽约时报》提到,他曾收到一份“班级中最好的论文”,段落简洁,举例恰当,论据严谨,令人生疑。随后这位学生承认,文章是 ChatGPT 写的。美国一家医疗保健创业公司的临床医生尝试让 ChatGPT 参加美国执业医师资格考试。这项标准化考试,考生至少要专门腾出 300~400 小时备考,最难的部分面向研究生。研究者下载了一份 2022 年 6 月的考题,确保 ChatGPT 从未训练过相关数据,也不知道考题答案,ChatGPT 考试成绩基本合格。该团队认为,这些结果表明,大型语言模型可能有助于医学教育,并可能有助于临床决策。

当地时间 2 月 10 日,微软创始人比尔·盖茨接受一家德国媒体采访时表示,ChatGPT 的重要性不亚于互联网的发明。而作为曾经的 ChatGPT 投资人、特斯拉 CEO 埃隆·马斯克也曾在推特发文称赞,“ChatGPT 好得吓人,我们离强大到危险的人工智能不远了。”

和其他聊天机器人相比,ChatGPT 显示出了更高的“情商”。用户会发现,ChatGPT 可以回答用户追问的问题,能承认错误,不断调整回答。问它《红楼梦》开篇“原来女娲氏炼石补天之时”的出处,它的回答是《山海经》,被提醒错误后,它很快道歉并调整答案:《封神榜》。它解释的理由中可以看到,它搜索的关键词是“女娲炼石补天”,再次追问,它说,女娲炼石补天的传说是一个历史悠久的故事,出现在多个文献和传统文化中,“因此,没有一个确定的出处”。

李笛说,ChatGPT 目前在三个地方有价值,它能帮用户生成一段内容作为初稿,例如邮件、短新闻等,用户再去修改,但现实中,“已有一些国外学生等用户,用它作弊,直接作为终稿提交”。

第二个价值在于,可以给用户提供“启发”和灵感。ChatGPT 可以把信息有条理地呈现出来,“假如你想找一个人讨论,尽管它的观点不一定准确,但它会为你提供一种思路。”李

笛说,第三个价值就是娱乐。除此之外,如果有人希望 ChatGPT 提供知识并对此深信不疑,“最好不要,没人能确保它的准确性”。

多位专家提醒,不必神话 ChatGPT 及其背后的大模型。“它可以像模像样写推荐信,但如果让它回答一些专业问题,你仔细看会发现,它可能是在一本正经胡说八道。”清华大学计算机系自然语言处理实验室副教授刘知远对《中国新闻周刊》说。

就连 ChatGPT 也承认自己的局限性。问及缺点,它回答:有可能会生成存在种族歧视、性别歧视等偏见的文本,因知识有限或不能理解语义而回答错误,不能处理复杂的逻辑和数学问题,而且,其拥有的知识只停留在 2021 年。新版必应和 Edge 浏览器则可以检索实时更新的新闻,甚至可以跟你聊过去一个小时发生的事情。

ChatGPT 发布仅一周,当地时间 2022 年 12 月 6 日,因用户大量搬运 ChatGPT 生成的错误答案,海外知名编程问答平台 Stack Overflow 暂时在该网站封杀 ChatGPT,认为“发布由 ChatGPT 创建的答案对网站及询问和寻找正确答案的用户来说,是非常有害的”。

“某种意义上有点像‘大力出奇迹’”

在李笛看来,就 ChatGPT 而言,“它不会对产业产生颠覆性影响,但大模型则会。”

2016 年前,小模型是人工智能理解人类语言的主流技术,下围棋或机器翻译等某一个具体任务,标注好数据,进入模型训练。但小模型无法用于其他任务,数据标注成本也极高。当时,人工智能领域催生了一个新的职业,人工智能数据标注员。“大家会嘲笑人工智能,说有多少人工,才有多少智能。”刘知远对《中国新闻周刊》说。

传统模式下,自然语言处理是一个非常严密的推理过程,不仅要识别每个词,还要处理词语间的序列,因此诞生了循环神经网络(RNN)模型。但 RNN 只考虑单词或者上下文信息,常常导致全文句意前后不连贯,或者理解复杂句子时出错。2016 年以前,当时的小冰、Siri 等问答系统或者机器翻译,都使用小模型。

“当时的人工智能行业,很多技术领域(比如机器学