新闻分析

# 人工智能现在有多“像人”？

新华社记者 彭茜 张漫子

随着人工智能大语言模型的不断迭代，其在类人智能方面取得了不少亮眼进展——会解读心理、察言观色，进行多轮实时语音对话，甚至还掌握了人类的欺骗、奉承等手段……这是否意味着我们距离实现通用人工智能已不再遥远？当人工智能“进化”得“更像人”，又会给人类社会带来哪些潜在风险？

五月二十三日，机器人在第七届数字峰会现场体验区表演太极拳。 新华社发

## 察言观色，接近真人

对他人心理状态进行解读的能力是人类社交的关键。近日发表在英国《自然·人类行为》杂志上的新研究发现，有的大语言模型在评估解读他人心理状态能力的测试中与真人表现相当，甚至在识别讽刺和暗示等测试项目中的表现胜过一些人。

解读和推断他人心理状态的能力被称为“心智理论”，是人类社交互动的核心能力，涉及沟通、共情和社会决策等。德国汉堡-埃彭多夫大学医学中心等机构研究人员测试了美国开放人工智能研究中心（OpenAI）发布的GPT系列大语言模型以及美国“元”公司发布的LLaMA2大语言模型在解读他人心理状态方面的表现，并与人类比较。

研究人员为大语言模型设置了通常用于评估“心智理论”涉及能力的5项测试：分别为识别错误信念、讽刺、失言、暗示和含有误导信息的奇怪故事，并将大语言模型的识别水平与1907名人类参与者相比较。研究发现，在所有5项测试中，GPT-4有3项测试（识别讽刺、暗示、奇怪故事）的表现优于人类，一项测试（识别错误信念）表现与人类相当，仅在识别失言的测试中逊于人类；而LLaMA2仅在识别失言的测试中胜于人类，其他测试项目表现均不及人类。

而OpenAI本月新发布的模型GPT-4o，则在与人的自然交互方面达到新高度，它可进行文本、音频和图像多模态的识别与回应，且更具“人情味”。它与使用者的语音对话基本无延迟，会倾听，能唠嗑，可变换各种语调。它还能识别人的面部表情、感知发言者的语气和情绪，并给出相应回应，让人惊呼“更像人”了。

## 欺骗人类，警惕风险

人工智能在拟人方面的进步不仅体现在“善解人意”上，甚至还学会了人类的欺骗、奉承等手段。此前，美国麻省理工学院等机构的研究团队在美国细胞出版社旗下《模式》杂志发表综述文章称，通过习得性欺骗，一些人工智能系统地学会了“操纵”他人。

研究发现最引人注目的欺骗案例是“元”公司的“西塞罗”人工智能系统，它被设计在一个虚拟外交战略游戏中作为人类玩家的对手。尽管“元”公司声称，该系统“在很大程度上是诚实和乐于助人的”，但研究人员发现，该系统在玩游戏时为达成比赛目标，背叛了盟友。

其他人工智能系统则具有在扑克游戏中虚张声势的能力，或在战略游戏“星际争霸2”中为击败对手而假装攻击，以及为了在谈判游戏中占上风而采取欺骗手段。

当人工智能掌握了欺骗技能，是否会给人类社会带来安全隐患？研究人员在文章中详述了人工智能欺骗带来的风险，如欺诈、制造假新闻、操纵选举等。

研究人员认为，“目前不可能训练出一个在所有可能的情况下，都不能实施欺骗的人工智能模型”，进而警示如果人工智能继续完善这套技能，人类可能会失去对它们的控制。因此建议尽可能用更多时间为未来人工智能产品和开源模型的更高级欺骗做好准备。

## 通用智能，尚未实现

尽管人工智能已在一些方面“进化”得十分像人，但相关专家指出，对大模型的“类人智能”需要有更清醒认知。目前距离实现完全类人、具备泛化能力的通用人工智能还有一定距离。

中国科学技术大学机器人实验室主任陈小平接受新华社记者采访时指出，要警惕人类对大模型产生“幻觉”。大模型学习大量历史数据，输出的表达方式符合许多人的语言习惯，让许多人误以为大模型会“说人话”或“理解人”，继而以为它具有某种“社会属性”，但实际上它没有心智。

“人工智能的内部工作原理与人类智能不同，但在某些局部是类似的。如果认为人工智能和人类智能相同，差别只在硬件载体的不同，就会做出很多不切实际的判断。”他说，目前对大模型测评的方法，基本上仍是传统软件的测评方法，因此需对这种方法得出的测评结果保持适度的审视态度。

汉堡-埃彭多夫大学医学中心的研究人员认为，大语言模型在“心智理论”涉及能力的测试中表现与人类相当，并非表明它们具有等同于人类的能力，也不意味着它们拥有人类“心智”。他们建议，未来研究可关注大语言模型在心理推理中的表现将如何影响人类个体在人机交互中的认知。

美国斯坦福大学计算机科学系教授李飞飞日前也在美国《时代》周刊刊文称，在通往通用智能的道路上，“感觉”是至关重要的一步，即拥有主观体验的能力。目前大模型并没有像人类一样的“感觉”，它可以说“自己脚趾痛”，尽管它根本就没有脚趾，它只是一个编码在硅芯片上的数学模型。

“我们还没有实现有感觉的人工智能，而更大的语言模型也无法实现这一目标。如果想在人工智能系统中重现这一现象，就需要更好理解感觉是如何在拥有实体的生物系统中产生的。”她说。

（新华社北京5月29日电）

新华视点

# APP自动续费的“坑”为何总填不平？

新华社记者 戴威 胡锐

移动互联网时代，手机APP成为人们生活中必不可少的工具。然而，一些APP自动续费屡遭质疑：从续费容易退订难，到默认或强制开通，再到自动续费悄然涨价、花的钱反而更多……

## 缘何问题反复出现、屡禁不止？

“当初同意自动续费，是为了消费更方便、更划算，没想到‘暗坑’这么多。”安徽合肥市民刘旭斌说。和刘旭斌一样，不少消费者感叹，APP自动续费容易退订难。

某知名投诉平台上，以“自动续费”为搜索条目的投诉件，已从2020年的3万多条增至如今超15万条。

“新华视点”记者调查发现，近年来，手机APP自动续费套路越发多样，尽管相关部门三令五申，部分APP经营者仍在自动续费设定里埋下“暗坑”。

2021年5月1日起施行的网络交易监督管理办法规定，网络交易经营者采取自动展期、自动续费等方式提供服务的，应当在消费者接受服务前和自动展期、自动续费等日期前五日，以显著方式提请消费者注意，由消费者自主选择；在服务期间内，应当为消费者提供显著、简便的随时取消或者变更的选项，并不得收取不合理费用。

“在自动续订、自动续费前，我压根没收到提示信息。”安徽芜湖市民郑成说，自己前不久在不知情的情况下，被某音乐APP默认自动续费。“虽然续费金额不高，但心里还是不舒服。”郑成说。

一些APP将取消自动续费的选项隐藏在难以察觉的位置，或通过繁琐步骤为用户取消自动付费设置障碍，被消费者形象地称为“套娃退订”。

在某外卖APP的一级操作界面，记者难以找到带有“自动续费”“续费管理”等字样的按钮。通过搜索相关教程，才在其“超级吃货卡”界面，经过四次点击关闭自动续费。记者发现，搜索“外卖APP如何关闭自动续费功能”，会跳出大量网友提问及相关教程。

一些APP“暗度陈仓”，不遵守原有低价合约悄悄涨价。

“2019年，我以每年197元的价格在某网盘APP开通自动续费会员。最近发现，从2020年开始，会员自动续费直接涨价到每年263元。而去年‘双十一’期间，在该APP官方网店购买年度会员仅需178元，远低于我的续费价格，这多少有些离谱。”一名消费者在某投诉平台上说。

记者在该平台以“自动续费涨价”为关键词检索发现，不少消费者反映，在APP上开通自动续费功能后，续订费用突然涨价，与当初签订的合约不符，违背了通过自动续费享受低价服务的初衷。

## 多次规范下为何屡禁不止？

套路陷阱困扰消费者久矣，有关部门多次出台政策法规加以约束，为何部分经营者依然我行我素？

——利益驱动大，监管难度高。消费者常用的手机APP如外卖软件、音乐软件等，涉及自动续费的金额通常在十余元至几十元间不等，单笔消费金额看似不高，但考虑到部分APP活跃用户在千万量级甚至破亿，自动续费带来的收益颇为可观。

以某音乐娱乐集团为例，财报显示其旗下APP在线音乐付费用户数量超过1亿。根据相关自动续费协议，这些音乐APP每月自动续费费用普遍在15元到25元之间；按最低15元计算，其会员续费月收入能达到十亿元量级。

此外，截至去年8月，国内市场上监测到的活跃APP数量达260万款。要对数量如此庞大的APP进行监管，对于相关部门来说颇有难度。

——违规成本低，打击力度弱。中国法学会消费者权益保护法研究会副秘书长陈音江说，APP自动续费乱象屡禁不止的一个重要原因是，经营者违法违规成本很低。消费者察觉到自己“被坑”后，诉诸法律手段维权的情况并不多见。

记者调查发现，和用户不计其数的线上投诉相比，也鲜有对APP经营者采取实质性处罚的案例。

今年3月，国务院公布《中华人民共和国消费者权益保护法实施条例》，自今年7月1日起施行，明确提及自动续费经营者应承担的责任，以及违反规定行为将受到的处罚。陈音江认为，这一条例从法规层面进一步规范自动续费，期待在具体执行时能真正“长出牙齿”。

——商家有空钻，用户维权难。消保专家认为，尽管相关政策法规提出“自动续费需征得用户同意”“不得默认勾选捆绑开通”“以显著方式提请消费者注意”等要求，但是相关规定还比较模糊。究竟如何才算显著、如何才算违规，尚不明确，这给部分经营者留下空子可钻。

全国政协委员、希肯国际文化集团董事长安庭说，此类纠纷涉及金额通常较小，消费者若想维权，花费的经济成本和时间成本较高。投诉、受理渠道不畅更加剧了维权困难。

## 多措并举维护消费者权益

对于消费者来说，开启自动续费，是为了节约交易成本、提高交易效率，但当前APP中存在的种种套路，却涉嫌侵犯消费者的知情权和财产安全权，让不少消费者直呼“上当”。受访专家建议，应加强监管力度、强化社会共治，维护消费者权益。刘旭斌表示，很多像他一样的消费者，常常苦于无处维权。他建议，相关部门应强化12315平台功能，畅通消费者投诉举报渠道，并督促相关APP经营者积极处理和回应消费者诉求，相关部门应对违法违规经营者严格调查处理。

此外，面对应用市场上海量的APP，全面监管难度较大，专家认为，可以通过大数据分析用户投诉、举报的信息，以此为监管方向和线索，积极监管、整治。

陈音江建议，应进一步完善相关法规，明确处罚措施，加大处罚力度，提高经营者违法违规成本。此外，相关政策规定应更为清晰，出台规范指引，避免部分经营者有漏洞可钻。安庭建议，对具有多次违法违规或其他严重违法违规情节的经营者，依法从严从重处罚。

安徽省消保委秘书处秘书长周玉战认为，自动续费产生的问题，往往具有点多、面广等特点，需加强社会共治，营造良好生态。他建议，通过媒体舆论监督、消保组织消费监督、行业组织强化自律、相关部门曝光典型案例等方式，群策群力推动问题源头治理。对于消费者来说，在使用自动续费功能时，也要擦亮双眼，仔细阅读相关条款，留存好相关证据。

（新华社合肥5月29日电）