

# AI时代,职业生态如何变化?

新华社记者 温竞华

## 如何让人工智能更『靠谱』

新华社记者 舒畅

越来越多人意识到,人工智能聊天机器人并不总是可靠。它可能答非所问、前言不搭后语,有时甚至“一本正经”地捏造不存在的信息。在1月下旬于新加坡举行的第40届人工智能促进协会年会上,如何让人工智能更“靠谱”、更“负责”,是这场国际人工智能学术会议的重要议题之一。

在研究者眼中,人工智能的缺陷对应着更细分的问题。比如,人工智能的置信度值是否与真实情况匹配等。一个来自意大利的研究团队在本届年会上提醒说,合理计算人工智能的置信度值非常重要。

在这项名为“盲目自信的陷阱”的研究中,参与者在人工智能协助下完成逻辑推理题。结果显示,当人工智能表现得非常自信时,参与者更容易采纳它的建议——但这些建议很可能是错的;然而,如果人工智能犹豫不决,用户却可能会拒绝正确的建议。

意大利米兰-比可卡大学的研究成员卡泰丽娜·弗雷戈西说,这两种情况都反映了现实问题:很多人工智能系统的置信度值没有被正确校准,从而向使用者传递了误导性的信号。

置信度值问题并非人工智能可靠与否的唯一体现。另一项来自印度和美国团队的研究,则从“认知疲劳”的角度,观察人工智能如何在不知不觉中变得不可靠。

美国南卡罗来纳大学人工智能研究所的研究成员里·玛尔瓦说,随着人类与聊天机器人对话时间越来越长,模型会开始“走神”:偏离原始指令、重复输出内容,甚至以自信的口吻生成不可靠的信息。研究团队将这种现象称为“认知疲劳”,且它并非偶发,而是一种系统性效应。

“认知疲劳是可以被发现、预测并干预的。”玛尔瓦说。团队设计了一个系统,通过追踪注意力衰减等三个关键指标,将人工智能何时开始“疲劳”可视化,并提供多种实时干预手段,如当人工智能开始偏离指令时,这套系统会重新输入原始指令,让对话回到“正确轨道”。

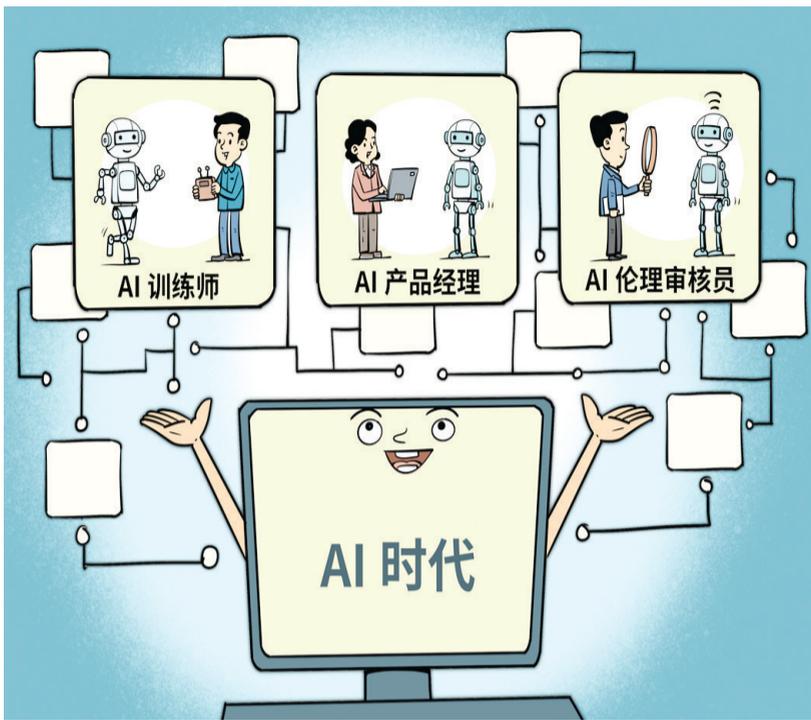
在大会现场,从单一的聊天机器人,到可以自主完成任务的智能体,再到多个智能体组成的协作网络,研究者们不只关注它们的能力与效率,也反复探讨它们的边界和可控性。

在不少参会者看来,完全放手让智能体自主完成任务,远没有想象中简单。

“目前,这些智能体的行动边界非常宽,它们能写代码,可以在计算机上做我们能做的任何事情……所以,我们应该把研发重点转向人类与智能体的协作,建立智能体与人类之间的透明交互,这具有现实紧迫性。”美国微软人工智能前沿实验室主任埃杰·卡马尔说。

微软首席科学官埃里克·霍维茨认为,当前种种不确定性,与人类尚未完全理解人工智能内部运作机制密切相关。

在历任主席对谈会上,他说:“今天的人工智能表现令人惊讶。我说‘惊讶’,是因为我们仍无法给出令人满意的科学解释:它们为何如此高效?如何失败?在哪些方面脆弱?对人工智能内部机制和原理的系统性科学探索,现在变得非常紧迫。我们必须打开这个‘黑箱’。”



新职业涌现 新华社发 朱慧卿 作

从AI训练师、AI产品经理、AI伦理审核员等新职业涌现,到“一人公司”等创业新范式兴起,人工智能正在加速重塑职业图景,也对劳动者的从业素质提出了新要求。

AI赋能下,技术范式的变革将创造哪些新就业新机遇?什么样的人能够更好地适应AI时代?

讯飞医疗科技股份有限公司内,AI医学研究员胡萍萍正在评估医疗大模型在复杂病例诊治中的表现,尝试发现模型在回应患者时缺乏“同理心”的细微问题。她和团队的目标,是让AI不仅“懂治病”,更要“懂人”。

胡萍萍曾是一名消化内科医生。5年前,她转型成为AI训练师,用医学专业知识参与训练医疗大模型,使其更好服务基层医生和居民健康管理。

“我们既是医学经验的传承者,也是技术应用的推动者。”胡萍萍说。

随着人工智能加速融入千行百业,从事数据标注的AI训练师、借助AI生成内容的动画师、统筹AI产品全流程的产品经理等新岗位持续出现,对复合型、应用型人才的需求不断扩大。

普华永道2025年全球AI就业晴雨表显示,全球几乎所有涉及AI应用的岗位数量都在增加;拥有AI技能的从业人员平均工资溢价达56%,是去年的两倍。58同城近期发布的报告也指出,平台已涌现近50类“人机协作”新岗位与40种智能新服务。

“人工智能技术正在深刻重构就业生态与职业图景,推动劳动者技能结构升级,向更高附加值方向演进,形成以人机协同、智能赋能为特征的就业新范式。”科大讯飞研究院院长刘聪说。

当前,具身智能、世界模型等技术方向加快突破,人工智能正从语言处理走向对物理世界的理解与建模。北京智源人工智能研究院发布的《2026十大AI技术趋势》认为,行业技术范式正发生深刻变化。

业内专家指出,AI正成为智能社会的重要基础设施,通过赋能传统行业、催生新业态,为提升就业质量提供新支点。

在机器人领域,具身智能推动人

形机器人向更高自主性和更强人机协同能力发展,逐步进入真实应用场景。

“人形机器人可广泛应用于工业、商业、文旅、医疗、家庭等各类场景。这一产业的规模化发展,将在设计研发、生产制造、检测认证、保养维修、场景应用等产业链各环节创造大量新就业岗位。”中国电子学会政策研究与国际合作处处长王桓说。

与此同时,AI也降低了创业门槛。“一人公司(OPC)”逐渐兴起——一个人借助AI工具即可完成内容生产、产品运营和服务交付。

从苏州提出打造OPC创业首选城市,到多地推出相关社区和支持政策,“单人+AI”的创业模式正加速走向主流。

“OPC正从个体创作转向具备规模化能力的‘超级个体’。”清华大学新闻学院、人工智能学院双聘教授沈阳认为,未来五年,这一形态有望成为数字经济的重要组成部分。

当AI可以在极短时间内生成大量内容,人的核心竞争力何在?

多位专家认为,在人机协作背景下,人类的想象力、判断力、审美能力、

批判性思维和情感互动能力,将成为不可替代的优势。

王桓指出,未来人才培养应强化交叉学科素养和综合能力,同时加快完善AI新职业培训体系,帮助劳动者顺利转型。

复旦大学计算与智能创新学院教授张军平作出一个判断:“斜杠”可能是未来人们主要的生活方式之一。

“因为我们无法预判AI会在哪个职业领域取得突破性进展,所以为提升抗风险能力,就要避免单一技能、单一职业依赖,提升综合能力。而且,AI的出现让人们可以快速地进入不同领域,大大缩短从新手到专家的过程。”张军平说。

张军平建议,青年一代要建立以“AI思维”为核心的认知范式,提升驾驭AI的能力以塑造发展新优势;要建立动态演进的知识体系,提升“跨学科融合+伦理引导”的未来竞争力。同时,警惕“因为过分依赖AI而变得更低能”的情况,尤其在青少年教育阶段不宜过早使用AI,应优先锻炼独立学习能力,培养独特的想象力和视角。

## 等你来赏! 本年度最大盈凸月29日现身夜空

1月29日将迎来本年度最大盈凸月。夜晚,感兴趣的公众可选择一个适合的时间段,欣赏一下这轮有些“发福”的月亮。

人们在地球上所看到的月球周而复始的圆缺变化现象称为月相,即月球的“相貌”,它每天都在变化,变化周期平均是29.53天。盈凸月是月相周期中从上弦月到满月之间的过渡阶段,通常出现在农历初九至十四之间,此时月球被太阳照亮的区域超过半球但未完全覆盖,形成西侧方向明亮的凸圆状外观。

中国天文学会会员、天津市天文学会理事杨婧介绍,盈凸月在日落前升起,次日日出前落下,黄昏时位于东

南天空,黎明前从西方地平线消失,因此在大半个夜晚都能观测到。

月球围绕地球公转的轨道是一个椭圆,在运转过程中,月球距离地球时远时近。所谓“年度最大盈凸月”,指的是月球在近地点附近时呈现的盈凸月视直径较大。

一些细心的公众发现,2025年12月曾接连出现过该年度第二大满月(5日)、最小残月(16日)和最大盈凸月(31日),而2026年1月就戏剧性地“复制粘贴”了这种情况,也出现了年度第二大满月(3日)、最小残月(15日)和最大盈凸月(29日)。这是巧合吗?

“这种情况实属巧合。每一个公历年,月球在围绕地球公转的过程中,

每月经过近地点和远地点,会呈现不同的月相,至于什么时候最大或最小,取决于月相周期、月球反射太阳光的面积、月球轨道偏心率、月球运行速度等多个因素。”杨婧说。

月相基本分为八个阶段:新月、蛾眉月、上弦月、盈凸月、满月、亏凸月、下弦月、残月。“这个过程中,随着月球的运动,我们在欣赏月球圆缺变化的同时,还会发现月球升起的时间、在天空中的位置也不一样。”杨婧说。

不同月相,各有其美。盈凸月如同一张拉满了九分的银弓,蓄势待发。如果天气晴好,喜欢赏月的朋友不妨走到户外去好好欣赏一番。

据新华社