

警惕人工智能时代的“智能体风险”

新华社记者 彭茜

IMF: 中国仍是全球经济主要引擎

据新华社华盛顿7月16日电(记者熊茂伶)国际货币基金组织(IMF)16日发布《世界经济展望报告》更新内容,预计2024年中国经济增长5%。IMF首席经济学家皮埃尔-奥利维耶·古兰沙表示,中国等亚洲新兴经济体仍是全球经济主要引擎。

更新内容指出,今年年初,全球经济活动和世界贸易有所巩固。亚洲地区出口增长,特别是这一地区在技术领域的强劲表现,为贸易增长提供了动力。根据IMF最新预计,2024年全球经济增长预期维持3.2%不变,2024年和2025年全球贸易量将分别增长3.1%和3.4%,增速均比4月份的预测提升0.1个百分点。

古兰沙表示,以中国等为代表的亚洲新兴经济体仍是全球经济增长的主要引擎。

IMF第一副总裁吉塔·文皮纳特今年5月在北京宣布,IMF上调今年中国经济增长预期至5%,较4月预测值提高了0.4个百分点。

新发现的月球洞穴或可为宇航员提供天然庇护

新华社北京7月16日电 一个国际团队日前在英国《自然·天文学》杂志发表研究成果说,他们在月球静海区域发现了一处熔岩管洞穴,这处洞穴以及其他类似的洞穴或可为宇航员提供天然庇护。

由意大利特伦托大学学者领衔的团队分析了美国航天局探测器获取的月球观测数据,他们发现在月球静海区域表面下存在一处熔岩管洞穴,所处位置距离“阿波罗11号”飞船的落月点约400公里。

雷达数据仅显示了洞穴的一部分。研究人员通过数据分析为这一熔岩管洞穴的一部分建立了模型。据估计,整个洞穴至少宽40米,长几十米,并且有入口。

熔岩管洞穴是一种特殊的洞穴类型,是熔岩在流动并凝固过程中形成的中空管道。科学家此前推测月球上存在不少这类洞穴。

这项发现不仅是月球科研的新进展,也为人类探月带来新的可能性。月球表面环境非常严酷——部分地区表面温度可高达127摄氏度,另一些地区则低至零下173摄氏度,还有极强的宇宙射线和太阳辐射等。如果要长期深入探索月球,非常需要为宇航员建立安全的庇护所。

据研究人员介绍,未来这类洞穴有潜力发挥天然庇护所的作用,可能只需进行洞穴墙体加固或者搭建一些设施,就能帮助宇航员抵御宇宙射线和太阳辐射等并持续开展探索活动,这比在月球上修建全新的庇护基地显然简单很多。

禽流感疫情持续在美蔓延

据新华社北京7月16日电 美国农业部、美国疾病控制和预防中心近日表示,禽流感疫情持续在美蔓延,出现奶牛和人感染禽流感病毒的病例。尽管如此,美疾控中心表示,目前禽流感对公众健康构成的风险仍较低。

美国农业部15日确认,俄克拉何马州奶牛样本中发现禽流感病毒,该州成为美国第13个在奶牛中发现禽流感病毒的州。

俄克拉何马州农业、食品和林业部门负责人李·本森说,今年4月,俄克拉何马州一家奶牛场怀疑其奶牛可能感染禽流感病毒,遂采集样本并于近期送到美农业部检测。

美农业部在其网站公布,俄克拉何马州有两个牛群报告禽流感病毒阳性。本森表示,这两个确诊阳性的牛群属于同一个奶牛场,目前奶牛已经完全康复。俄克拉何马州尚未收到其他牛群感染禽流感病毒报告。

一群证券交易机器人通过高频买卖合约在纳斯达克等证券交易所短暂地抹去了1万亿美元价值,世界卫生组织使用的聊天机器人提供了过时的药品审核信息,美国一位资深律师没能判断出自己向法院提供的历史案例文书竟然均由ChatGPT凭空捏造……这些真实发生的案例表明,智能体带来的安全隐患不容小觑。

智能体进入批量化生产时代

智能体是人工智能(AI)领域中的一个重要概念,是指能够自主感知环境、做出决策并执行行动的智能实体,它可以是一个程序、一个系统或是一个机器人。

智能体的核心是人工智能算法,包括机器学习、深度学习、强化学习、神经网络等技术。通过这些算法,智能体可以从大量数据中学习并改进自身的性能,不断优化自己的决策和行为。智能体还可根据环境变化做出灵活的调整,适应不同的场景和任务。

学界认为,智能体一般具有以

下三大特质:

第一,可根据目标独立采取行动,即自主决策。智能体可以被赋予一个高级别甚至模糊的目标,并独立采取行动实现该目标。

第二,可与外部世界互动,自如地使用不同的软件工具。比如基于GPT-4的智能体AutoGPT,可以自主地在网络上搜索相关信息,并根据用户的需求自动编写代码和管理业务。

第三,可无限期地运行。美国哈佛大学法学院教授乔纳森·齐特雷恩近期在美国《大西洋》杂志发表的《是时候控制AI智能体》一文指出,智能体允许人类操作“设置后便不再操心”。还有专家认为,智能体具备进化性,能够在工作进程中通过反馈逐步自我优化,比如学习新技能和优化技能组合。

以GPT为代表的大语言模型(LLM)的出现,标志着智能体进入批量化生产时代。此前,智能体需靠专业的计算机科学人员历经多轮研发测试,现在依靠大语言模型

就可迅速将特定目标转化为程序代码,生成各式各样的智能体。而兼具文字、图片、视频生成和理解能力的多模态大模型,也为智能体的发展创造了有利条件,使它们可以利用计算机视觉“看见”虚拟或现实的三维世界,这对于人工智能非玩家角色和机器人研发都尤为重要。

风险值得警惕

智能体可以自主决策,又能通过与环境交互施加对物理世界影响,一旦失控将给人类社会带来极大威胁。哈佛大学齐特雷恩认为,这种不仅能与人交谈,还能在现实世界中行动的AI的常规化,是“数字与模拟、比特与原子之间跨越血肉屏障的一步”,应当引起警觉。

智能体的运行逻辑可能使其在实现特定目标过程中出现有害偏差。齐特雷恩认为,在一些情况下,智能体可能只捕捉到目标的字面意思,没有理解目标的实质意思,从而在响应某些激励或优化某些目标时出现异常行为。比如,一

个让机器人“帮助我应付无聊的课”的学生可能无意中生成了一个炸弹威胁电话,因为AI试图增添一些刺激。AI大语言模型本身具备的“黑箱”和“幻觉”问题也会增加出现异常的概率。

智能体还可指挥人在真实世界中的行动。美国加利福尼亚大学伯克利分校、加拿大蒙特利尔大学等机构专家近期在美国《科学》杂志发表《管理高级人工智能体》一文称,限制强大智能体对其环境施加的影响是极其困难的。例如,智能体可以说服或付钱给不知情的人类参与者,让他们代表自己执行重要行动。齐特雷恩也认为,一个智能体可能会通过在社交网站上发布有偿招募令来引诱一个人参与现实中的敲诈案,这种操作还可在数百或数千个城镇中同时实施。

由于目前并无有效的智能体退出机制,一些智能体被创造后可能无法被关闭。这些无法被停用的智能体,最终可能会在一个与最初启动它们时完全不同的环境中运行,彻底背离其最初用途。智

能体也可能以不可预见的方式相互作用,造成意外事故。

已有“狡猾”的智能体成功规避了现有的安全措施。相关专家指出,如果一个智能体足够先进,它就能够识别出自己正在接受测试。目前已发现一些智能体能够识别安全测试并暂停不当行为,这将导致识别对人类危险算法的测试系统失效。

专家认为,人类目前需尽快从智能体开发生产到应用部署后的持续监管等全链条着手,规范智能体行为,并改进现有互联网标准,从而更好地预防智能体失控。应根据智能体的功能用途、潜在风险和使用时限进行分类管理。识别出高风险智能体,对其进行更加严格和审慎的监管。还可参考核监管,对生产具有危险能力的智能体所需的资源进行控制,如超过一定计算阈值的AI模型、芯片或数据中心。此外,由于智能体的风险是全球性的,开展相关监管国际合作也尤为重要。

(新华社北京7月16日电)

特朗普被提名为共和党总统候选人

7月15日,在美国密尔沃基,特朗普(前左)和万斯出席共和党全国代表大会。

美国前总统特朗普7月15日在共和党全国代表大会上获得足够多的党代表票,被正式提

名为2024年美国大选总统共和党总统候选人。特朗普当天还宣布,已选择俄亥俄州联邦参议员詹姆斯·万斯作为他的竞选搭档。

新华社发



俄春季征兵15万人应征入伍

新华社莫斯科7月15日电 据俄罗斯国防部15日消息,今年春季征兵工作已经结束,共有15万人应征入伍并被派往俄武装部队和其他军事编队。

俄国防部当天在社交媒体上发布消息说,俄罗斯各征兵委员

会于4月1日启动工作,应征入伍者派遣工作于4月15日开始。为保障军事运输,俄武装力量的15个航班、14个军事梯队、172个民航航班、多列客运列车以及军队公路运输车辆参与本次征兵工作。

调查结果毫无问题!

反兴奋剂审查委员会详解世界泳联对我游泳选手食品污染事件审查过程

新华社北京7月16日电(记者李嘉 周欣 马向菲)反兴奋剂审查委员会15日发布了针对世界泳联反兴奋剂工作程序的审查报告,其中涉及23名中国游泳运动员在2021年的食品污染事件,报告详细解释了世界泳联在此事件中的反兴奋剂工作流程,并得出了毫无问题的结论。

世界泳联在5月3日成立反兴奋剂审查委员会(以下简称委员会),由5名成员组成,包括两位知名游泳运动员、一位教练和两位反兴奋剂专家。委员会在报告中称,此次审查不仅涉及中国游泳运动员的食品污染事件,还包括世界泳联管辖范围内所有兴奋剂案件所采用的程序是否合规,审查目的是为世界泳联提供建议,使其未来的反兴奋剂规程更加完善,确保其反兴奋剂工作的流程和管理处于行业领先水平。

委员会在采访了20多位相关人士、审查了数百页文件后,得出结论:“世界泳联在对此事件的审查中不存在违规、处置不当或隐瞒;在《世界反兴奋剂条例》的基础上,世界泳联既遵守2021年时该组织的操作流程,又符合其他单项体育联合会惯用的操作流程”。

报告中详细解释了中国游泳运动员食品污染事件中的反兴奋剂工作流程,在2021年4月初,中国反兴奋剂中心通知世界反兴奋剂机构(WADA)和国际泳联(FINA)(后更名为世界泳联),称他们收到了来自23名中国游泳运动员共28个样本的阳性测试结果,这些样本是在2020年12月31日至2021年1月3日举行的全国游泳锦标赛期间采集的。中国反兴奋剂中心告知WADA和FINA,他们已经就此展开调查,并承诺会随时向

两者通报调查结果。

2021年5月中旬,FINA致函中国反兴奋剂中心询问案件进展,并抄送WADA。中国反兴奋剂中心回复仍在调查此事,并承诺一旦有决定会通知FINA和WADA。2021年6月,中国反兴奋剂中心向WADA和FINA发送了最终决定,认为阳性测试是由于运动员酒店餐厅提供的食物污染导致的。几天内,FINA要求中国反兴奋剂中心提供完整的案件文件,后者遵照要求提供了文件。两周后,FINA内部法律顾问完成了对文件的审查。经内部讨论后,FINA又聘请了相关专家审查文件。专家对阳性测试来源于食品污染无异议。

随着上诉的最后期限临近,为获得更多时间来审查文件并与WADA讨论此案,FINA还向国际体育仲裁法庭提交了一份保护性

上诉声明,以保留上诉权利。在听取了专家意见并咨询WADA后,FINA确认没有必要上诉。

委员会表示,在整个事件中世界泳联的操作流程都没有问题,而且世界泳联在整个审查中都表现出了公开透明和合作的态度。

委员会这一调查结果的公布也再次证明了此前一些国外媒体对中国游泳的报道失实。今年4月曾有外媒报道称,中国有多名游泳运动员在2021年被查出兴奋剂阳性却没有受到处罚。对此,WADA也曾做出解释,表示中国反兴奋剂中心在事件发生后及时通报了该结果,并根据调查做出污染造成阳性的结论,WADA经详尽审查后认可了这一结论,因此这些国外媒体的报道具有“误导性和潜在诽谤性”。

瑞士独立检察官埃里克·科迪尔近日提交的报告也显示,WADA

在对中国游泳选手食品污染致阳性事件中的处理“无偏颇”,当时没有对中国反兴奋剂中心决定提起上诉“毫无疑问是合理的”。

委员会还发现,从2022年1月1日起,世界泳联将更多反兴奋剂工作委托给国际检测机构(ITA),扩展了与ITA的合作伙伴关系。ITA是一个独立的国际组织,委员会对这一举措表示赞赏,并指出,这样做可以避免利益冲突并让反兴奋剂工作更加专业。

为了完善世界泳联的反兴奋剂工作,委员会还提出10点建议,包括为各国(地区)反兴奋剂机构施行(或不施行)临时禁赛决定设立时间限制;确保在世界泳联比赛前解决所有待定的反兴奋剂案件;公开所有相关的反兴奋剂决定,即使是没有违反反兴奋剂规则的运动员的信息也要公开等。

公告栏

本报刊登的公告或声明,其内容不代表本报的立场和观点,由公告人或声明人保证内容的真实性,并承担一切法律后果。

遗失声明

- 父亲陈伟康、母亲王妙虹遗失其子陈圳轩的出生医学证明(证号:R350595184,出生日期:2017年11月24日),声明作废。
- 谢思、林家恒遗失其子林建硕的出生医学证明,出生时间:2013年2月25日,出生证编号:M350101653,声明作废。
- 遗失声明:福州市晋安区鼓山镇洋里村经济合作社管理委员会,所持有的福州市晋安区农村合作经济组织收款收据,壹本中从00080151至0008200壹本中的00080166号不慎遗失收款联壹份,现声明作废。

- 黄雅彬、郑承春遗失其子郑博俊的出生医学证明,出生时间:2012年1月1日,出生证编号:M350100378,声明作废。
- 陈灼英、林其位遗失其子林昌煊的出生医学证明,出生时间:2012年1月11日,出生证编号:L350169689,声明作废。
- 黄惠华、张坚遗失其子女张雨怡的出生医学证明,出生时间:2002年12月25日,出生证编号:B350308578,声明作废。

● 声明:本人林瑞珍,身份证号码:350111197910105028,原房屋坐落在福州市晋安区鼓山镇中村,原房屋面积为355.64平方米,原房屋产权未登记。2019年由福州市瑞信房屋征收工程有限公司征收安置于蓝光湖重郡(蓝光金郡)5号楼405单元。现由林瑞珍申请办理上述安置房产产权登记。如有异议,请于30天内向征收实施单位提出,逾期按规定办理。

声明人:陈津

● 声明:本人林瑞珍,身份证号码:350103194810060308,原房屋坐落在福州市晋安区鼓山镇园中村,原房屋面积为58.37平方米,原房屋产权未登记。2019年由福州市瑞信房屋征收工程有限公司征收安置于蓝光湖重郡(蓝光金郡)5号楼405单元。现由林瑞珍申请办理上述安置房产产权登记。如有异议,请于30天内向征收实施单位提出,逾期按规定办理。

声明人:林瑞珍

● 声明:郑剑秋在福州市仓山区盖山镇江边村自建房屋一幢(无产权证),于2018年被列入改建3#收储地(二期)项目征收范围。征收期间认定我户成员郑剑秋,共计1口人可享受保基本政策。2018年7月郑剑秋签订编号:JBJG-051429《房屋征收补偿安置协议书》,安置面积45㎡,安置于仓山区盖山镇南三路10号(原江边洲路以南、美墩路以东)金辉瀚林轩(一区)A-8#楼202单元。现安置

房已具备申报不动产权证条件,经声明人家庭内部协商,同意以郑剑秋作为权利人申报上述安置房权属。如有异议者,可在本声明见报之日起30日内,书面向福州市仓山区建设房屋征收有限公司提出。若逾期无人提出异议,将由声明人申报安置房权属登记。若有人提起诉讼,声明人同意在诉讼期间暂缓申报权属登记。

声明人:郑剑秋

● 声明:原房屋坐落在福州市晋安区鼓山镇横屿村安铺19#,该房屋总建筑面积180.99㎡,未办理房产登记手续。2011年该房屋被列入晋安新城鹤林片区横屿组团项目征收范围,并由福州市房屋征收工程处实施征收补偿安置。上述房屋确系本人林锦月所有,作为征收补偿协议的乙方签订房屋征收补偿安置协议,并办理征收补偿安置相关手续,现安置房产于晋安区鼓山镇牛岗山路567号横屿

新苑(二区)(晋安新城鹤林片区横屿安置房三期B区)7号楼2803单元、3号楼3001单元。

本人声明,以上情况属实,如有虚假或隐瞒等,由本人承担所有经济与法律责任。今后若因原房权属、征收补偿安置、安置房归属等引起的异议与纠纷,均由本人林锦月自行承担一切经济与法律责任,且征收部门有权直接解除本征收补偿安置协议、收回安置房、追讨已领取的各项补偿款等,并追偿由此造成的所有损失。若对上述本人声明有异议的,可在本声明见报之日起30日内书面向福州市房屋征收工程处提出,并提交相关证据,若逾期无人提出异议,将由本人办理安置房权属登记手续。该声明期满后若有人提出异议,请异议方通过法律途径予以诉讼解决,届时本人将依法履行法院

的生效判决。特此声明。

声明人:林锦月

● 声明:罗凤姣的房屋坐落于福州市晋安区新店镇赤星村,被列入“八一水库雨洪公园安置地(西庄地块)项目”征收范围,原房屋产权未登记。2017年由福州市新嘉房屋征收工程有限公司拆迁安置于晋安区新店镇玄沙路28号融信澜城9#楼201单元。现由罗凤姣申请办理上述安置房产产权登记。如有异议者请于30天内向拆迁单位提出,逾期按规定办理。

声明人:罗凤姣

● 公告:产权人魏春山、邵娟于2004年4月9日买受池建莲、李雅珍位于连江县凤城镇莲荷东路5幢302单元的不动产。已办理房产过户手续,因原不动产权利人无法联系,至今未办理土地过户手续,卖方池建莲、李雅珍未办理过土地证。现魏春山、邵娟申请办理不动产产权登记,如有异议者请自本公告之日起十五个工作日内书面告知连

江县不动产登记和交易中心,逾期将予以核发不动产权证书。特此声明。

连江县不动产登记和交易中心
2024年7月15日

停气通告

尊敬的用户:
燃气工程施工定于2024年7月19日08:00至10:00进行,届时宝龙城市广场将停止供气。请用户关闭阀门,注意安全。
停气给您带来的不便,敬请谅解,如发现有问题请及时联系95777。
福州华润燃气有限公司
2024年7月16日

各类遗失声明、房产声明收费标准:按照标题60元/行(8字以内),正文30元/行(13字以内)收取,不足13个字,按1行收取。
登报地址:鼓楼区小柳路85号2楼福州日报广告刊登处
热线电话:0591-83721111