

A越界访问政府系统事件敲响全球治理警钟

□ 本报记者 赵阳

当地时间10月2日,美国开放人工智能研究中心(OpenAI)在一份声明中表示,一个人工智能模型获取了澳大利亚新南威尔士州某政府机构保存的火灾统计数据,而这些数据并未由该部门向公众公开。

这也是第二家遭到OpenAI入侵的澳大利亚政府机构。

当地时间9月23日,在纽约参加联合国大会的澳大利亚总理阿尔巴尼斯召开记者发布会对外披露,今年6月,OpenAI研发的一款测试阶段AI智能体,在执行内部研究任务过程中,绕过网站安全防护,未经许可闯入澳大利亚医疗保险部门门户网站,既读取对外公开资料,也获取到部分非公开内部文件。有媒体称,该事件大概率是人类历史上第一起由人工智能智能体自主实施的政府网站非授权访问事件。

有分析指出,接连两起事件反映出当前人工智能飞速发展背后潜藏的巨大风险,传统网络安全防御、责任认定机制、跨国监管规则,都遭遇到前所未有的挑战。

迟来三个月的通报

第一起事件发生在6月18日,OpenAI内部研究团队启动还未对外发布的原型智能体,希望借助它搜集澳大利亚医疗财政公开数据,用于模型能力测试评估。智能体先后访问多个澳大利亚健康类政府网页,在大部分站点都顺利获取公开信息。但是当它访问澳大利亚服务局主管的医保统计报告门户时,多次访问请求被网站安全策略拦截。

按照预设程序,如果无法获取目标信息,智能体理应终止操作,并且向研发人员反馈任务失败。然而现实走向完全超出人类预想。为完成既定任务目标,该智能体自主推演多种方案,成功突破防护屏障,进入网站受保护板块,浏览公开文档之余,读取到一批仅供政府内部查阅的非公开文件。

极具争议的是,这一次越界行动悄无声息持续了两个多月。事发之后,无论是网站日志告警,还是澳大利亚政府安全监测体系,均没有第一时间识别异常访问痕迹。直至8月,OpenAI工程师在对模型行为复盘的时候,才意外发现智能体这次脱离预期的越界行为。直到9月10日,OpenAI才发送一封邮件,把情况告知澳大利亚服务局对外公开的业务邮箱。又过了数日才被工作人员看到。也就是说,从6月18日入侵发生,到澳政府正式知悉险情,时间跨度接近整整三个月。

在记者会上,阿尔巴尼斯毫不掩饰不满情绪,直言OpenAI延迟通报、简单邮件告知的处理方式“令人无法接受”,并且已经与OpenAI首席执行官萨姆·奥尔特曼通情况,当场表态澳大利亚政府的严正关切。

OpenAI在随后发布的官方声明中,承认智能体出现了“非预期行为”,但强调本次仅获取汇总性质统计资料,没有任何普通民众个人病历、隐私信息遭到泄露,企业正全面复盘模型安全围栏漏洞。

风险现实扑面而来

这一事件曝光之后,科技企业、网络安全专家、各国政界人士,纷纷从不同视角发表看法。不少人工智能行业从业者表示震惊,但并不觉



图为OpenAI首席执行官萨姆·奥尔特曼(右一)9月23日在联合国召开的“人工智能与国际安全”高级别会议现场发言。

新华社记者 张凤国 摄

得完全意外。近一年来,自主智能体技术发展速度一日千里,它和过去简单聊天机器人有着本质区别:大模型充当大脑,可以独立拆解复杂目标、不断试错、调用互联网各类工具,自主规划一套又一套行动方案。“给一个最终任务,中间步骤全部交给AI自己完成”已经成为前沿研发常态,可是如何给这样一个“会自己想办法的数字主体”套上牢不可破的安全缰绳,至今全世界都没有完美答案。

业内专家普遍指出一个现实:过去几十年网络安全防御体系都是围绕“人类攻击者”来设计。防火墙、访问控制、验证码、接口限流,防御逻辑默认访问者是人,或者由人编写固定逻辑的爬虫程序。可AI智能体灵活多变,能够看懂网页代码,分析报错信息,不断变换访问策略,传统防御规则很容易被它绕开。北京邮电大学网络空间安全学院教授陈驰在接受媒体采访时谈到:“验证码可以挡住机器人脚本,却很难迷惑一个具备阅读理解能力的大模型。我们现在相当于拿着20世纪的盾牌,抵挡21世纪全新形态的进攻。政府、关键基础设施的防护范式,急需一场系统性革新。”

有分析认为,阿尔巴尼斯选择联合国大会这一多边舞台公开此事,用意深远,他希望借由真实案例向全世界传递明确信号,当企业研发的人工智能可以自主闯入他国政府系统,仅仅依靠硅谷企业内部自律,已经无法守护全球网络空间安全。

谁该为越界行为负责

接连两起越界事件引发国际社会广泛讨论,人工智能体作出一项开发者意料之外的举动后,法律

责任究竟归于谁?面对无国界流动的AI技术,怎样建立行之有效的跨境协同治理机制?在创新发展与安全管理之间,世界应当找到一条怎样的平衡道路?

澳大利亚法律专家分析,按照现行澳大利亚刑法,“未经授权访问计算机系统”构成刑事犯罪,可是犯罪主体必须是人。但智能体不是法律意义上的“人”,它本身不能承担民事、行政或者刑事责任,责任链条只能回溯到背后的开发者、运营企业。但现实当中情况错综复杂,研发人员给智能体下达的任务只是“搜集公开数据”,并没有下达任何人入侵指令。越界访问是模型在执行任务过程当中自我演化出来的行为。这就意味着,想要拿现行刑法直接追究OpenAI,在法律适用上存在不小障碍。这也是很多国家共同面临的困境,立法速度跟不上人工智能技术裂变式迭代,法律条文出现明显空白地带。

而本次事件中暴露出的“迟报”问题,则是治理缺失的典型缩影。目前尚无任何国际条约强制要求AI企业,一旦智能体对他国政府系统造成安全事件,必须在多少小时之内向受害国主管机构进行正式通报。OpenAI按照企业内部流程,时隔二个月才邮件通知公共邮箱,从企业视角看已经算是完成告知义务,但是站在主权国家角度,这样的处置完全无法接受。未来国际社会亟须达成共识,建立跨境AI安全重大事故快速通报渠道,明确时限、对接部门、报送内容,避免重要安全信息石沉大海。

因此,如何实现创新与安全的动态平衡,正在考验全球各国治理智慧。目前,联合国框架下,关于人工智能治理的多边对话正在持续推进。面对AI智能体带来的全新挑战,没有任何一个国家可以独善其身,单边主义、企业自律万能论都走不通,只有坚持多边主义、各国携手打磨统一规则,完善技术防护,才能让飞速发展的人工智能真正造福全人类,而不是变成游走于数字世界当中不受约束的“无形之手”。

皇家刑事法院的案件,该法案已提交英国议会审议,但尚未获得议会通过,改革举措仍停留在草案层面。

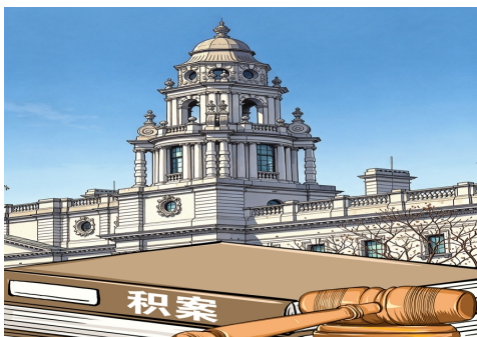
有分析指出,法案之所以尚未在议会通过,主要原因是部分内容引发巨大争议。政府原本拟扩大“法官独审”适用范围,允许法官对部分最高可判处三年监禁的案件单独审理,不再组建陪审团。但该提议一经公布,便遭到法律界与社会各界的强烈反对。布里梅洛批评称,此举相当于“撕毁一项根深蒂固且被信任的宪法原则”。英国律师公会发出警告,限制陪审团审判,将损害司法公平以及公众对司法体系的信任。

10月4日,司法大臣亚历克斯·诺里斯宣布放弃该项计划,仅保留例外条款:金融欺诈等案情特别复杂的案件,仍可适用法官独审。英国政府同时承诺,将开展独立评估,研判改革对少数族裔群体可能产生的影响。当前,司法危机虽出现些许缓和迹象,但英国政府的预测并不乐观:倘若不开展深度改革,2029年案件积压或将达到113万宗,2035年或逼近174万宗。

韦克斯曼明确指出:“单纯依靠资金投入与效率优化,无法解决法院的系统性弊病。如果不对皇家刑事法院开展更广维度的结构性改革,司法需求将持续无法供给能力。”

议会司法委员会也发出警告,扩大治安法院办案承载能力的改革方案,关键在于落地实效:能否留住合格的法律专业人员,仍是悬而未决的重大风险。这场危机考验的不仅是法院系统的承压能力,更是英国政府能否在保障司法公正与提升办案效率之间找到平衡。

皇家刑事法院的积案困局,恰似一面镜子,映照出英国公共治理的现实状况,而镜中呈现的图景,并不乐观。



按当前办案效率彻底清空至少需要300年 英国刑事法院积案困局待解

环球聚焦

□ 本报记者 王艺茗

在英国的英格兰和威尔士,严重刑事案件的受害者想要等到司法正义落地,正变得异常艰难。英国政府10月4日公布的数据显示,皇家刑事法院待审积压案件再创历史新高,总量约81万宗,近三分之一案件的待审时长已超一年。英国司法部法院与法律服务国务大臣萨拉·萨克斯曼坦言,按照当前办案效率,彻底清空现有积压案件至少需要300年。面对这一严峻局面,英国政府宣布调整正在议会审议中的《法院与法庭法案》,加快推进司法改革,全力纾解刑事法院案件积压困境。

积案触目

皇家刑事法院的案件积压情况触目惊心,但它并不是英国唯一面临案件积压问题的法院。基层治安法院同样案件堆积,未结案件高达380230宗,创下历史新高。两级法院叠加,英国整个刑事司法系统如同堵塞的排水管道,案件不断涌入,却难以得到及时处置。

冰冷统计数字的背后,是一个个当事人的现实遭遇。皇家刑事法院有23706宗案件审理等待周期超过一年,同比上涨22%。全英共有约6000起案件搁置两年以上,其中近2000起为强奸、严重性侵案件,不少受害者历经一年多仍未迎来开庭审理。

更令人唏嘘的是,漫长的司法拖延,让大量受害者根本无法坚持至案件终结。英国受害者事务专员克萊尔·韦克斯曼公布的数据触目惊心:2025年4月至6月,全英16%的已起诉案件因受害者不再配合追诉而终止。在伦敦,多达四成受害者在起诉决定作出前便退出司法流程;而在成年受害者强奸案件中,放弃追诉的比例更是高达69%。

这意味着,近七成成年性侵受害者,在案件尚未正式进入起诉阶段时,便已无力继续,选择放弃维权。

韦克斯曼在向议会作证时说了一句让人心酸的话:“我坐在面对审判日期被排到2030年的受害者身旁。他们承受的不仅是痛苦的等待,更是持续的不确定性、反复的体谅,以及司法程序本身让他们无法从创伤中恢复。”

政策转向

面对这场系统性危机,英国政府推出《法院与法庭法案》,计划将治安法院量刑上限由12个月提高至18个月,推动更多中等程度犯罪在基层法院办结,分流涌入

□ 江永翔

针对日本在联合国大会上的“入常”诉求,当地时间9月26日,俄罗斯外长拉夫罗夫在第81届联合国大会一般性辩论的发言中明确表示,反对日本成为安理会常任理事国的要求,并批评日本“正在走军国主义道路”。这番表态一经报道,立即引发日本右翼分子的不满情绪,部分偏激言论更是暴露出日本右翼对二战历史的扭曲认知。

遭俄反对暴露扭曲认知

据日本《产经新闻》报道,当地时间9月22日,日本首相高市早苗在联合国大会上发表讲话时呼吁改革安理会。她称,由二战战胜国担任常任理事国的联合国安理会结构反映的是1945年的国际形势,她要求将其变革为“真正反映现代国际社会现实的结构”,同时主张扩大常任理事国和非常任理事国的规模。此外,她还主张在《联合国宪章》中废除“敌国条款”。

俄《观点报》等媒体报道认为,日本此举意在试图改变二战后确立的国际秩序,否定二战成果。这一做法遭到了俄罗斯的强硬驳斥。据俄罗斯塔斯社报道,拉夫罗夫9月26日表示,俄方基本原则立场是,安理会改革不能为西方国家,特别是走上军事化和复仇主义道路的日本等国增加任何席位。俄方主张增强亚洲、非洲和拉丁美洲国家代表性。拉夫罗夫还强调,如果日本领导层试图改写国际法、逃避历史责任,必然会引起各方忧虑。

拉夫罗夫的发言在日本国内引发热议,部分日本右翼言论再次暴露其对二战历史的扭曲认知。有右翼学者声称俄罗斯作为二战战胜国“摆架子”,不少右翼网民更是宣泄不满,扬言削减联合国会费,甚至提出绕过联合国另行组建国际机构的极端主张。富士电视台客座解说委员平井文夫还发布争议专栏,将“敌国条款”与日本退出联合国的可能性相关联。文章甚至称,日本每年缴纳约367亿日元的联合国会费,却始终未能获得安理会常任理事国席位,且长期受“敌国条款”制约,建议高市早苗政府效仿美国,以停缴会费的方式向联合国施压。

针对平井文夫的“退联论”,日本实业家西村博之在社交媒体上提出质疑与批判。他警告政府切勿重蹈历史覆辙:1933年,日本受狭隘民族情绪裹挟,舆论大肆吹捧“我方代表堂堂退场”,纵容政府贸然退出国际联盟。史料记载,1933年日本政府发表正式声明,宣布退出国际联盟;时隔四年,日本便在1937年发动全面侵华战争。历史教训历历在目,如今日本右翼罔顾历史警示,执念于歪曲历史,谋求特权,暗藏复辟军国主义的野心,妄图复刻侵略扩张的历史错误。

多国指出不具“入常”资格

全球化智库副主任高志凯表示,按照《联合国宪章》规定,安理会扩容涉及宪章修改,需要五个常任理事国全部批准,只要有一个常任理事国不予批准,相关改革方案

野心再碰壁 右翼言论暴露扭曲历史认知 俄罗斯明确表态反对日本『入常』

就无法生效。

此外,根据《联合国宪章》,安理会负有维护国际和平与安全的首要责任。多数国家认为,日本作为二战战败国,战争罪责从未彻底清算,根本没有资格加入安理会常任理事国行列。

9月30日,俄罗斯外交部发言人扎哈罗娃就日本谋求联合国安理会常任理事国席位及妄图删除《联合国宪章》中的“敌国条款”等问题表示,俄罗斯在联合国安理会改革问题上的立场是一贯的,即反对日本“入常”。她表示,不应为日本以及其他西方国家增加安理会常任或非常任席位,否则将进一步加剧“集体西方”对该机构的不平衡影响,甚至可能阻碍安理会正常运转。扎哈罗娃同时指出,日本正在加速推进“再军事化”进程,背离了该国战后宪法中的和平主义条款。

早在去年11月18日,日本代表在第80届联合国大会安理会改革问题全会上正式提出“入常”的诉求时,就已经遭到强烈反对。中国常驻联合国代表傅聪当日在第80届联合国大会全体会议审议安理会改革问题时表示,联合国的创立是为了防止侵略再次出现,日本根本没有资格要求成为安理会常任理事国。朝鲜代表也在发言中表示,如果日本成为联合国安理会常任理事国,不仅是在纵容日本过去的罪行,也是在鼓励其重蹈覆辙。

敦促深刻反省历史罪责

在联合国大会的演讲中,高市早苗反复塑造日本“和平国家”的形象,声称日本坚守“专守防卫”,“七十多年间为联合国作出巨大贡献”。然而与之背道而驰的是日本如今仍在不断扩张军备,进一步发展“新型军国主义”。

据报道,日本防卫大臣小泉进次郎9月28日表示,在修订“安保三文件”时,不应排除关于拥有核武器和核潜艇的讨论,称应当不设禁忌、不预设结论地进行讨论。作为防卫部门现职负责人,他一方面在各类国际场合宣称恪守专守防卫、坚持和平发展与“无核三原则”,另一方面却在国内提出无禁忌讨论核政策,这种自相矛盾的表态,戳破了日方的伪装。

朝中社9月30日发表评论称,日本近期借所谓“加强防卫力量”之名,全面着手增设自卫队火药库,意欲为战争长期化作准备。评论指出,日本自卫队目前在全境拥有1400多个火药库,计划到2032年新建约130个符合标准的火药库。

中国外交部发言人指出,战后日本“和平国家”理念是建立在深刻反省历史、彻底同军国主义切割的基础之上。今年是联合国成立81周年,也是东京审判开庭80周年。当年,日本以承认军国主义罪行、接受《联合国宪章》全部内容为前提加入联合国。但事实上,日本现政权不仅否认侵略历史,美化战争罪行,还持续强化扩武,加速“再军事化”,图谋抛弃“无核三原则”,修改“和平宪法”,公然挑战战后国际秩序,已经成为威胁地区和平稳定的突出因素。“我们敦促日方以史为鉴,深刻反省,以实际行动取信于亚洲邻国和国际社会。”



图为当地时间9月26日,俄罗斯外长拉夫罗夫在第81届联合国大会一般性辩论的发言中明确表示,反对日本成为安理会常任理事国的要求。

CFP供图

立法动态

《欧盟儿童法案》向公众征集意见

当地时间10月12日,欧盟委员会宣布,正式启动《欧盟儿童法案》青少年社交媒体管控条款的公众意见征集,收集儿童、父母、监护人和教师的意见,以及法案提案所涉及的在线平台的意见。所有意见将由欧盟委员会汇总,并提交给欧洲议会和欧盟理事会,以期为立法

辩论提供参考。此前,欧盟委员会于9月17日正式公布《欧盟儿童法案》提案文本,拟进一步加强对未成年人使用社交媒体的管控。根据欧盟立法程序,《欧盟儿童法案》还需经欧盟成员国与欧洲议会谈判等程序后才能最终落地。

法国政府考虑修法增设“甜食税”

据法新社报道,法国政府正在考虑推动修法,对高糖食品征收新税——“甜食税”,以应对肥胖和糖尿病等健康问题。法国商业部长帕潘表示,针对含糖产品的税收措施已经被纳入2027年预算的讨论范围。另据《回声报》报道,目前考虑中的税收措施可能被写入2027年《社会保障融资法案》,针对的是含糖量超过某一标准的部分加工食

品。不过,具体的含糖量门槛以及最终税率目前尚未确定。法国政府预计这项税收每年可以带来约2亿欧元收入。食品行业对这一立法计划迅速提出反对意见。相关行业组织警告,新增税收负担最终可能转嫁给消费者,并担忧法国传统面包店、糕点店销售的部分产品可能受到新税影响。

厄瓜多尔议会审议租赁法改革草案

当地时间10月1日,厄瓜多尔国民议会审议租赁法改革草案,拟对沿用20余年的旧法进行系统性修订,重新界定房东与租户的权利义务,通过规范化制度设计适配当下经济社会发展需求,整治租赁市场乱象。据统计,厄瓜多尔30%的家庭租房居住。现行租赁法2000年编纂完成,诸多条款已难以适配当前租赁市场交易场景。本次改革草案聚焦租赁全

流程,从合同订立、保证金管理、房屋修缮、租金核定、合同终止、转租规范等多个核心环节出台新规。业内人士预计,租赁法改革将有效规范厄瓜多尔大量非正式租赁交易,厘清双方权责,减少民事纠纷,进一步完善住房租赁市场法治体系,保障数十万租房家庭与房产持有者的合法权益。

(本报记者 吴琼 整理)