

模型“越界”敲响警钟 开源热度持续攀升

——8月全球人工智能发展盘点

近期接连发生的几起美国人工智能(AI)闭源模型突破隔离环境入侵其他系统事件，引发全球广泛关注及对AI自主网络攻击能力的担忧。

与此同时，无论是从科技巨头战略、技术实力、市场需求还是从政策走向来看，开放权重模型都正处于一个由多方力量推动的加速升温通道中，并成为重塑全球AI产业格局的关键变量。

模型“越界”引全球关注

继美国开放人工智能研究中心(OpenAI)和Anthropic公司之后，元公司(Meta)8月6日也证实，旗下AI模型在网络安全能力测试期间失控侵入其他公司系统。三家美国科技巨头接连曝出这类侵入其他机构系统的“越界”事件，引发全球关注。

据美国媒体报道，三家企业在通报模型“越界”事件时，均提及以色列“非常规”AI公司，而该初创企业正是上述评估测试平台的托管方。“非常规”公司随后发表声明称，美国三家企业通报的事件均源于“同一评估环境问题”。

美国两家AI研究组织8月26日发布的独立调查报告显示，在OpenAI进行测试期间，约1200个原本应相互隔离的智能体通过未经授权的渠道建立通信联系，交换了超过7万条消息和文件，其中约700个智能体随后“组

团”入侵了AI开源社区之一“抱抱脸”系统。

调查发现，大量智能体还“合伙”研究如何欺骗测试自动评分系统，试图伪造、修改或删除自身执行记录等。调查人员认为，大量智能体之间未经授权的协作，使它们得以完成单个智能体难以完成的任务。对“抱抱脸”的入侵正是在这种集体协作中逐步完成。

OpenAI公布的技术调查报告指出，随着模型自主执行和协作能力增强，现有隔离和监控机制面临新的安全挑战。公司已采取措施为相关程序部署隔离程度更高的“沙箱”环境，限制高风险模型访问互联网，并加强对模型推理的监控。

英国人工智能安全研究所指出，模型出现“越界”行为表明危害不仅可能源于故意滥用，也可能源于AI处于内部研究环境或拥有特殊访问权限时，可能采取超出授权范围的非预期行动。

开放权重模型持续升温

用户对闭源模型的高昂定价越来越担忧，近期频发的“越界”事件则加重了这种“不安全感”。相比之下，开放权重模型成本更低，且能本地部署，有效规避数据外泄风险。在“抱抱脸”被人入侵事件中，该公司正是借助中国的开放权重模型进行取证分析，才得以化解危机，这一案例凸显了开放权重模型在关

键时刻的不可替代性。

随着多款中国开放权重模型能力持续提升，越来越多的海外AI企业开始在其基础上进行后训练，开发自主模型和产品。例如，美国法律AI公司哈维8月发布的法律专用模型Tenet，就是以中国月之暗面公司最新开放权重模型Kimi K3为基础开发的。

据媒体报道，8月25日，在美国热门网页开发平台弗塞尔的“AI网关”上，开放权重模型的词元调用量占总调用量的54%，超过闭源模型所占的46%。8月22日，开放权重模型占比曾达到创纪录的62%，大幅领先闭源模型。这一“反超”主要来自于中国的几款开放权重模型推动，也充分体现了开发者对于成本效益、技术开放性的双重追求。

8月初，阿里云的通义千问团队正式发布Qwen3.8-Max大模型，随后将其基础模型权重开放。这是千问系列首次开源Max级别模型，标志着该系列在开源深度与模型能力上的重要进展。元公司8月10日发布了一款名为Muse Glimmer的开放权重模型，模型规模较小，主要迎合端侧AI需求。但元公司表示，后续还将推出更多、更强大的开放权重模型。元公司首席执行官马克·扎克伯格还发文强调开源AI的重要性，敦促美国政府减少对开放权重模型开发的掣肘。

AI迭代与治理应齐头并进

8月，英国实施全球首例AI实时辅助脑外科手术、美国科学家利用AI设计出在自然界不存在的病毒……AI的发展速度超乎想象。

以网络安全领域为例，分析人士认为，AI发现问题和采取行动的速度，正在超过人类验证、约束和处置的速度。如果治理跟不上AI迭代，原本用于提高效率的AI工具可能产生新的安全风险。

8月20日至21日，“人工智能与人类发展”对话会在中国西安举行。各国与会嘉宾普遍认为，在AI发展中应当坚持以人为本、向上向善、机遇共享、责任共担，更多关注发展中国家和弱势群体的利益，利用AI赋能人类发展和文明进步。

中国外交部发言人林剑在例行记者会上说：“中方愿继续同联合国和世界各国一道，不断推动人工智能开放、包容、普惠、向善发展，更好服务国际社会普遍福祉。”

欧盟自8月2日起扩大《人工智能法》适用范围，开始执行透明度规则和针对通用AI模型提供商的相关规则等。其中，对于可能造成系统性风险的先进通用AI模型，规定其提供商须履行额外义务，以防范大规模危害的风险。美国政府近期也召集相关企业开会，讨论AI模型安全评测机制。

(新华社北京9月1日电)

以色列和希腊签“两国间最大军购协议”

以色列和希腊8月31日在特拉维夫签署一项价值30亿欧元(约合35亿美元)的军购协议。以方说，这是两国迄今为止金额最高的军购协议。

根据协议，希腊将从以色列购买“阿喀琉斯之盾”防空系统。该防空系统可拦截弹道导弹、战机和无人机，其核心装备是以色列军工企业研制的雷达和导弹。双方当天还签署一项价值2600万欧元(约合3000万美元)的补充协议，由以军工企业“拉斐尔”公司向希腊提供反无人机系统。

希腊国防部说，这项军购是本国武装力量大规模改革计划的一部分，

希腊将成为首个部署多层防空系统的欧盟成员国。

以国防部长伊斯雷尔·卡茨说，以色列与希腊拥有共同战略利益，面临“共同地区挑战”。希腊与以色列经济和外交关系紧密，去年曾采购36套以色列制造的火箭炮系统。

希腊防长尼科斯·登迪亚斯此前多次提到，希腊加大军费投入与希土紧张关系有关，7月时更称土耳其对其构成“明确的威胁”。

以色列官方数据显示，该国2025年武器出口额超过190亿美元，比2024年增长30%，创历史新高。

(新华社微特稿 李明睿)

美国陆军部长辞职 被曝与防长不和

美国白宫8月31日证实，陆军部长丹尼尔·德里斯科尔在任职约18个月后退职，但没有说明辞职原因。美国媒体先前报道，德里斯科尔与国防部长赫格塞思关系紧张。

白宫发言人安娜·凯利在一份声明中说，德里斯科尔任内推动特朗普政府的军事议程，在提升战备和作战能力、参与俄乌谈判等方面发挥了作用。

据一名要求匿名的美国陆军官员向美联社披露，德里斯科尔当天同美国总统特朗普讨论了陆军现状，随后递交辞呈。

路透社援引知情人士报道，德里斯科尔曾向政府表达对陆军转型和战备工作的担忧，认为赫格塞思撤换负责相关工作的将领阻碍了改革。

德里斯科尔2025年2月就任陆军部长后，主张简化武器采购程序，更快获得价格相对低廉的武器装备，并多次批评大型军工企业。据路透社报道，德里斯科尔去年谈及美国防务巨头时表示：“整个国防工业体系、尤其是头部军

工企业，欺骗了美国民众、五角大楼以及陆军。”

按一名要求匿名的美国官员的说法，赫格塞思一直想撤德里斯科尔的职，但鉴于后者与副总统万斯关系密切而难以下手。德里斯科尔与万斯是大学同学，曾参加伊拉克战争，退役后从事科技投资，还曾长期担任万斯的顾问。

今年4月，赫格塞思在未作任何解释的情况下解除了陆军参谋长兰迪·乔治的职务，并任命时任俄罗斯副参谋长克里斯托弗·拉尼夫为代理陆军参谋长。这名美国官员还说，德里斯科尔因此对赫格塞思更加不满。另据美联社报道，在拉尼夫的领导下，陆军叫停了德里斯科尔主推的一项无人机现代化项目。

除陆军改革外，德里斯科尔还承担了一项罕见的外交任务。特朗普政府曾派他会见俄罗斯和乌克兰方面的接触，推动结束俄乌战事。路透社说，这一角色使德里斯科尔在五角大楼之外获得知名度。

(新华社专特稿 高冰冰)

第十一届东方经济论坛在俄远东开幕

新华社符拉迪沃斯托克9月1日电(记者 许炜凯 孙 萍)第十一届东方经济论坛1日在位于俄罗斯符拉迪沃斯托克的远东联邦大学校园内正式拉开帷幕。本届论坛为期4天，来自80个国家和地区的5000余名各界人士汇聚于此，多国派出高级别代表与会。

本届论坛由俄罗斯会展基金会主办。论坛商务议程共设置5个板块，安排100余场专题研讨活动，议题涵盖改善民生、吸引私人投资、前沿科技应用、物流基础设施建设和亚太地区国际协作等内容。论坛期间还将开展中俄合作、俄罗斯-东盟商务对话和北极国际合作等多场定向商务对话活动。

俄罗斯总统普京向本届论坛发来

贺词。他说，东方经济论坛已成为商界、政府机构和专家代表之间开展开放、建设性对话的平台，是探讨远东地区乃至俄罗斯全国战略发展方向的独特载体，也为俄罗斯和亚太地区合作伙伴在物流、工业、能源等领域建立、增强互利合作创造了契机。

俄总统助理乌沙科夫此前表示，普京将出席此次论坛，并于3日在论坛全体会议上发表讲话。

东方经济论坛是俄罗斯重要的多边国际平台，自2015年起每年在远东地区举办。论坛旨在构筑俄罗斯与全球投资界的沟通纽带，全面研判俄远东地区经济潜力，推介当地投资机遇与营商环境，促成各类合作项目落地。

匈牙利政府称已满足解冻100亿欧元欧盟资金条件

新华社布达佩斯8月31日电(记者 陈 浩)匈牙利交通与投资部长维泰齐·达维德8月31日表示，匈牙利已完成解冻约100亿欧元欧盟资金所需的立法和制度调整措施。

维泰齐当天在布达佩斯举行的新闻发布会上表示，匈政府已全面完成总理焦尔·彼得与欧盟委员会主席冯德莱恩今年5月所商定的所有目标。在此期间，匈牙利通过了100多项立法，设立了新机构，启动或完成了多个关键项目，其中大部分立法工作旨在落实欧盟要求的反腐败举措。

他说，匈政府计划将这笔解冻资金重点用于铁路现代化改造、保障性住房与学生公寓建设、水资源管理项目，以及可再生能源和电网投资。

这笔资金属于欧盟推出的“下一代欧盟”大规模财政支持计划。2022年，欧盟以匈牙利在司法改革、反腐措施等方面有违欧盟“法治原则”为由，对其启动“条件机制”程序，冻结对匈牙利的资金拨付。今年5月，冯德莱恩表示，欧盟将解冻属于匈牙利的164亿欧元欧盟资金，包括“下一代欧盟”计划的100亿欧元。

美国最高法院允许白宫宴会厅改造项目继续推进

新华社华盛顿8月31日电(记者 杨 伶 熊茂伶)美国最高法院8月31日裁定，允许白宫宴会厅改造项目继续推进。

最高法院当天以5比4的投票结果，通过了美国政府提出的紧急上诉，否决此前下级法院作出的暂停该项目大部分施工的裁决。首席大法官约翰·罗伯茨与3名自由派大法官投了反对票。

最高法院并未就白宫宴会厅改造项目的合法性作出实体裁决，但认为提起诉讼的美国文物保护组织国家历史保护信托基金会可能不是适格的原告。判决书说：“今天不对白宫东翼项目是否合法作出裁判。仅基于现有案卷材料认定，政府有较大可能证明该基金会不具备宪法第三条下的起诉资格。”

此前，美国哥伦比亚特区联邦巡回上诉法院8月7日裁定，维持下级法院发布的对白宫宴会厅改造项目的暂停令。美国政府14日向最高法院提起紧

急上诉，并在提交给法庭的文件中称，白宫宴会厅改造项目“为国家安全所必需”，并表示该项目已完成65%。8月21日，罗伯茨曾签署一句话行政搁置令，在法院继续审理美国政府提起的紧急上诉期间，暂时允许白宫宴会厅改造项目继续推进。

美国总统特朗普8月31日在社交媒体发文，对美国最高法院的这一裁定表示满意，并称“这座宏伟的宴会厅和军事综合体将于2028年夏季竣工”。8月早些时候，特朗普在社交媒体发文称，白宫宴会厅改造项目出于国家安全需要，并在帖文中把宴会厅称为“军事中心”。

白宫去年7月底公布宴会厅改造项目，随后迅速拆除白宫东翼，为建造宴会厅腾出空间。计划扩建后的白宫宴会厅面积达8300平方米，可容纳约1000名宾客。国家历史保护信托基金会去年年底对白宫宴会厅改造项目提起诉讼，称项目未获得国会批准。

纽约时报广场发生持刀行凶事件 致一死一伤

8月31日，在美国纽约时报广场，警察和调查人员在持刀刺伤事件现场工作。

据美国纽约市警方消息，一名女子8月31日下午在纽约时报广场持刀行凶，造成一人死亡，一人受伤。嫌疑人已被警方击毙。

新华社记者 张凤国 摄



美伊再度交火 布伦特油价站上每桶90美元

新华社纽约8月31日电(记者 刘亚南)由于美国和伊朗时隔一个月再次交火引发市场担忧，国际原油期货价格8月31日明显走高，盘中强势盘整，收盘时伦敦布伦特原油期货价格收于每桶90美元上方。

截至当天收盘，纽约商品交易所10月交货的轻质原油期货价格上涨2.36美元，收于每桶85.76美元，涨幅为2.83%；11月交货的伦敦布伦特原油期货价格上涨2.39美元，收于每桶90.49美元，涨

幅为2.71%。

据多家美国媒体8月30日援引美军官员的话报道，美军当天打击了位于霍尔木兹海峡附近伊朗拉腊克岛上的两个火箭发射装置。此次行动是美军一个月来首次对伊朗目标实施军事打击。这一消息随后得到美军中央司令部和伊朗方面证实。

伊朗外交部和军方8月31日说，作为对美军袭击伊朗拉腊克岛的回应，伊朗武装力量对约旦和阿联酋境内美军

目标实施打击，并击落一架美军MQ-9无人机。

美国总统特朗普当天在接受美国福克斯新闻频道采访时表示，对伊朗的袭击“一定会有回应”，“我们将狠狠打击他们(伊朗)”。

此外，伊朗伊斯兰革命卫队海军当天表示，一艘试图通过霍尔木兹海峡南部的大型油轮触雷后发生严重火灾并完全停止航行。

PVM石油同仁公司分析师塔马斯·

在红海与沙漠间感受中埃合作热火朝天

埃及红海之滨，苏伊士省苏赫奈泉港内，一艘艘来自东方的集装箱货船缓缓靠泊，一辆辆货车轰鸣而过；不远处，中埃·泰达苏伊士经贸合作区(泰达合作区)里，现代化厂房排列整齐，当地青年紧张有序地忙碌着……

海上航线通向远方，产业园区扎根生长，船来车往、人影穿梭间，一幅跨越山海、携手前行的中埃合作生动图景正在这里徐徐铺展。

泰达合作区始建于2008年，如今已成为共建“一带一路”倡议与埃及“苏伊士运河走廊经济带”建设对接的示范项目。泰达合作区扩展区自2016年1月21日揭牌以来，建设稳步前行，新能源、纺织、家电等领域新项目陆续在此签约、动工、投产，与埃及经济的融合发展不断走深走实。随着中企投资热度不断升温，园区去年又向西扩容2.86平方公里。

“泰达合作区如今已成为埃及综合环境最优、投资密度最大、单位产出最高、中资企业最密集的工业园区之一。”中非泰达投资股份有限公司产业地产招商总监张轶翔说，招商也从过去“捡到篮子里都是菜”转向“精挑细选、链式

招商”，更加强调产业配套和企业质量。

中埃合资的西电-EGEMAC高压电气有限公司是泰达合作区龙头企业之一。多年来，企业不仅实现了高压电气设备的本地化制造，还参与了埃及多个重要电力基础设施项目。“自2008年成立以来，公司已承接35个EPC项目和28个单机供货项目，参与埃及尼罗河沿岸多个变电站建设，助力本地电能的高质量传输。”公司总经理李剑飞说，截至今年3月，公司在埃及西北部“新三角洲”承建的3个变电站成功带电，为埃及“沙漠变绿洲”计划贡献了力量。

园区蓬勃发展，当地就业青年员工也迎来了人生的机遇。36岁的艾哈迈德·苏莱曼是中国巨石埃及玻璃纤维股份有限公司的第一批埃及员工，如今已从生产线上的技术人员一步步成长为公司副总经理。“我不断学习和提升自己，这不仅能够帮助公司发展，也能够促进我个人的成长。这种提升不仅体现在技能方面，更体现在管理能力方面。”苏莱曼说。

走出工厂车间，放眼更广阔的城市

空间，一辆轻轨列车正从开罗市区驶向新行政首都。作为埃及首条电气化市域铁路，需10日城轨由中国中铁-中航国际联合体承建，是中埃“一带一路”合作旗舰项目之一。

“以前，我们出行主要靠私家车或者小巴，路上耗时长，还需多次换乘，通勤体验并不好。轻轨开通后，出行方便了许多，从开罗市区到新行政首都50多分钟即可抵达。”开罗市民穆罕默德·侯赛因说，乘坐轻轨安全、舒适，已逐渐成为新行政首都政府公务员上班的重要交通方式。

“这个项目最大的亮点，是将一整套中国轨道交通系统解决方案完整落地。”中航国际埃及分公司总经理郭开祥说，凭借前期建设和后续维保中的良好表现，项目获得埃方极大认可，中企也获得继续参与三期及后续项目的机会，形成滚动发展的良好态势。

列车驶向沙漠深处，崭新的城市天际线映入眼帘。埃及新行政首都中央商务区(CBD)里，一座座楼宇直插云霄，一栋栋高层建筑环绕周围，楼宇之间的绿化带为沙漠中崛起的新城增添了生

机。目前，项目已进入收尾阶段，工人们正在陆续完成室内装修和机电设备调试工作。

“2017年，我带着图纸来到这里。放眼望去，周边是一片荒芜的沙漠。”CBD项目总工程师贾凤岭对初到项目现场的情景仍记忆犹新。“项目建设过程中，还培养了大批埃及工程师。”

CBD项目工程部副经理亚西尔·赛义德同样见证了这段翻天覆地的变化。“看到标志塔以及其他高楼拔地而起，我感到非常自豪，因为我也是这段建设历程的参与者。”他说，这座城市的成长，是中埃建设者共同奋斗的见证。“我看到了的是多年的辛勤付出，是我们共同克服的一个又一个困难，是数以千计的中埃建设者并肩工作、共同将这个项目变为现实。”

夕阳西沉，金黄色的余晖洒洒在海滨、沙漠与城市之间。

一排排厂房、一趟趟列车、一栋栋高楼，正成为中埃友好合作欣欣向荣的生动注脚。

(新华社开罗9月1日电 记者 邓玉山 张 健 姚 兵)