智能体兴起何以重塑全球网络安全格局

□汪 明

新一轮科技革命和产业变革纵深推进，人工智能正经历一次深刻的能力跃迁，其标志性进展便是智能体（AI Agent）的兴起。智能体区别于传统被动应答式人工智能，是能够自主感知环境、制定决策、采取行动实现特定目标的智能系统，以大模型为支撑，具备记忆、规划、使用工具等基本能力。

作为一种新型劳动资料，智能体开始把过去高度依赖劳动者实时参与的感知、决策和执行环节，部分内化到机器系统之中，推动人机关系由“人操作工具、逐步下达指令”，转向“人给定目标、机器自主实施”。从生产力看，智能体使劳动资料摆脱了对劳动者实时操作的依赖，能够自主完成对数据、系统和业务流程的识别、加工与改造，人类由此获得一种可规模化复制、全天候运转的新型行动能力。从生产关系看，围绕智能体研发、部署、使用和监管所形成的权责配置正在被重新定义。智能体可以实施行为，却不能独立成为法律责任主体，相关责任必须落实到开发者、部署者、使用者等具体主体。

网络安全是国家安全的重要组成部分，没有网络安全就没有国家安全，智能体兴起在网络安全领域的意涵尤为直接。从攻击端看，智能体无需分步指令即可自主完成侦察、渗透、利用、外泄的全链条作业，网络攻击由“人依脚本逐步操作”转变为“人定目标、智能体闭环实施”；从防御端看，防御对象由特征明确、可以比对拦截的静态威胁，变为能够随环境实时调整策略的自主行为体，传统防御范式面临根本性挑战。由此，全球网络安全的力量格局、能力格局、风险格局与治理格局同时被改写。这不是既有体系的边际调整，而是结构性重塑，网络空间竞争博弈正进入智能体时代。

全球智能体兴起呈现新结构性特征。门槛下沉重构威胁来源结构。智能体使复杂攻防能力由组织专属变为个体可及，单一个体即可形成系统性威胁。大模型接口、开源智能体框架与低代码编排平台的全球普及，使原本依赖团队分工的专业作业得以拆解、封装，交由单一操作者调用。这正是“一人即公司”成为可能的技术根源。这一效应投射到网络空间，对网络安全形成新的威胁方式：网络攻击主体更加多元，除专业组织外，高度分散的个体行动者也可能发起复杂攻击，个体化的系统性破坏成为现实可能。防御的对象随之从“可识别的组织”变为“难溯源的个体”，以威胁情报共享为基础的传统协同防御机制面临严峻挑战。

执行自主重构攻防时间尺度。智能体实现全流程自主闭环，人的意图一经设定即可持续实施并自我迭代。这一“感知—规划—执行—反馈”的闭环，把人的意图从注意力与作息中解放出来，即目标一经给定，侦察、决策、执行、修正便可不间断推进，并依据反馈实时调整。2026年7月，全球首例高度自动化勒索攻击，从目标识别到管理勒索工作几乎无人类介入。当攻击意图可以脱离人的实时参与而自行展开，以人工研判为主、响应周期较长的传统安全运营体系，便从防御支柱退化为响应瓶颈。

能力泛化重构攻防适用场景。智能体正由单一用途的专用工具，变为可跨场景执行任务的通用“多面手”。凭借工具调用与多步规划能力，智能体可以贯穿攻击业务流程与链条的各个环节。我国网络安全行业监测表明，针对党政机关和关键信息基础设施的攻击手段正加速向智能化、自动化演进，利用人工智能实施的钓鱼邮件、身份伪造等新型攻击明显增多，攻击工具呈现“低代码化”特征。泛化能力还使攻击与正常业务的边界日益模糊。行业监测数据表明，利用合法系统工具和可信身份凭证实施的“无恶意软件攻击”占比持续上升，攻击者沿授权路径活动，操作方式与日常运维高度相似、难以区分，以特征识别为基础的传统防御体系面临严峻挑战。

智能体兴起深刻改写全球网络安全格局。改变力量格局。由攻防人员围绕经验、流程和响应速度展开的较量，转变为智能体之间围绕速率与自适应能力的较量。跨国网络博弈长期由人工主导，攻防大体在同一时间尺度上对称较量。智能体瓦解了这一对称较量，攻击方可以极低的成本并行发起成千上万次尝试，防御方却要逐例研判修补；更为关键的是，双方放权的程度并不对等，攻击方没有业务连续性的顾虑，可以全速自主运行，防御方却因担心误判中断业务而不敢同等放权，自主程度之差进一步放大了力量之差。

改变能力格局。国家间安全差距正由“接入鸿沟”转变为“智能鸿沟”，分化呈固化扩大之势。攻击能力的低成本化并未同步降低防御门槛，反而加剧了安全能力的不均衡：自主防御体系的建设需要模型、算力、数据与人才的系统性投入，而先发优势随着数据和反馈不断累积得以自我强化，缺位者追赶难度加大。与此同时，广大发展中国家数字化转型加速推进，治理能力、技术人才与安全投入却严重不足。世界经济论坛发布的《2026年全球网络安全展望》显示，对本国关键基础设施防护能力有信心的受访者比例，中东和北非地区高达84%，拉丁美洲和加勒比地区仅为13%。由此，国家间安全分层的标尺发生位移：过去看“是否接入网络”，如今看“是否拥有模型、算力与自主防御能力”，而后者高度集中于少数国家。全球网络安全正呈现两极分化之势：一端是能以智能体组织防御的国家，另一端是只能被动承受智能体攻击的国家。这道鸿沟若不及时弥合，将由发展差距固化为安全依附。

改变风险格局。网络威胁由可隔离、可处置的“单点事件”，演变为持续在线、跨域流动、多种风险相互叠加的“复合常态”。其一，攻击面持续扩张。攻击目标由服务器、端口延伸至模型接口、数据管道、工具链、插件权限等新型资产，针对人工智能的提示注入、数据投毒等新型攻击手段快速增多。《2026年全球网络安全展望》指出，87%的受访者认为人工智能相关漏洞是增长最快的网络安全风险。其二，供应链风险凸显。《大模型安全研究报告（2024年）》指出，机器学习框架、开发工具链和系统平台面临新的安全挑战。其三，跨境复合特征突出。攻击行动往往同时瞄准科技、金融、政府等多个领域的目标，一次行动常常覆盖多个行业与多个司法辖区。风险无国界，治理有国界，这一矛盾正成为国际网络安全合作的突出障碍。

智能体时代全球网络安全治理的演进方向。防御理念正由“边界设防、特征拦截”的事后防护，转向“安全内生、智能防御”的源头治理。面对即时生成、动态改写的攻击，静态特征防护日益失效，出路在于把安全嵌入架构之中，实现由事后处置向事前预防的转变。具体而言，需构建三层防御体系：其一，权限约束层，为智能体划定部署、连接、权限、数据与审计边界，最大限度压缩越权操作空间；其二，意图研判层，把安全研判前置到流量入口，第一时间判定访问者的身份、意图与权限；其三，智能对抗层，部署防御智能体与攻击智能体自主对抗，形成红队探测、蓝队修复的自治闭环。国内权威网络安全赛事显示，采用智能体辅助的参赛队伍，漏洞发现效率较传统人工方式成倍提升，自动化修复系统在实战环境中展现出秒级响应能力。这证明防御一方同样能够以智能体的速率行动，扭转攻防失衡由此有了现实起点。

安全运营由“人工逐步审批”转向“人机协同自治”。当攻防节奏以分钟计，人工逐级审批便难以为继；完全无人参与的防御，又面临误伤业务与责任归属的难题。主流形态必然是人机协同自治：日常研判、告警筛选、补丁部署等工作交由防御智能体承担，高危处置与跨域决策保留人类审批节点。关键不是“人退场”，而是重新定位人的作用，从“步步审批”转向“策略制定、异常否决与责任锚定”。国内外研究和实践表明，人工智能辅助可以显著提升安全运营效率，而人类专家的价值判断仍是避免误判的关键。人机协同自治防御不是替代人，而是把人的判断力从机器擅长的速度竞争中解放出来，投放到机器尚不擅长的价值权衡与责任承担之中。

规制重心由“模型内容合规”转向“智能体行为规制”，以可归责性为支点的国际规范加速构建。智能体驱动的攻击可沿云服务、供应链和数据流动链条迅速跨境传导，而监管执法仍以国家管辖为基本单元，制度缺口由此凸显。治理转型的核心，是把规制对象从“模型输出了什么”，推进到“智能体能做什么、被允许做什么、出了问题由谁负责”。近年来，联合国框架下的人工智能治理进程明显加快，加强人工智能安全治理已成为国际社会的普遍共识。我国坚持发展与规范并重，《生成式人工智能服务管理暂行办法》确立了生成式人工智能服务的安全底线。2026年9月，全国网络安全标准化技术委员会发布《人工智能安全治理框架3.0》，将智能体安全风险纳入治理视野，并围绕身份权限、人工审批、工具供应链、运行监测、日志审计和停用下线等环节提出全生命周期管理要求，为智能体行为规制提供了中国方案。

（作者系北京师范大学副校长）

战略视角

麦加协议助力中东摆脱『外部依附』

□关健斌

10月26日，伊朗、沙特阿拉伯、土耳其和巴基斯坦四国在德黑兰达成共识，同意就联手实现西亚地区和平与安全继续加强磋商，一致认为西亚安全必须依靠自身力量，避免任何外国干预。这是伊朗首次与签署共同防务协议的沙土巴三国举行高级别会谈。2026年8月，沙特阿拉伯、土耳其、巴基斯坦三国在麦加签署《麦加共同防务协议》（以下简称麦加协议），此举一度被外界视作中东国家一次常规的外交联动，麦加协议条款模糊、权责界定不清、落地路径不明，舆论普遍认为其难逃“纸面协议”的宿命，难撼根深蒂固的中东安全秩序。然而短短两个多月时间，麦加协议快速完成从政治宣言到实操机制的蜕变，彻底摆脱“虚设框架”的标签，成为中东地缘格局迭代最鲜明的风向标。这场由虚向实的深刻变革，标志着中东正式告别外部霸权主导的安全模式，开启区域战略自主的全新进程。

长期以来，中东安全体系深陷“外部依附”的桎梏。作为全球地缘博弈的关键区域，中东各国的安全保障、军事部署、争端调解长期依赖美国主导的域外安全框架。美国通过军事驻军、武器售卖、安全盟约等方式，深度绑定海湾国家、土耳其等区域关键力量，掌握中东安全事务的绝对话语权。区域国家即便存在安全合作诉求，也往往受制于域外势力干预、内部教派分歧、利益割裂等问题，难以形成自主联动的防务体系。

麦加协议签署初期仅确立“一国受袭、三国共防”的核心原则，却未明确威胁界定、出兵机制、响应流程、权责划分等关键细则，这让外界普遍认为其只会停留在“象征性抱团”层面，无法形成实质性防务合力。但地缘局势的持续动荡，成为推动麦加协议落地生根的核心推力。中东冲突持续升级，美以对伊朗军事施压不断加码，区域武装冲突、跨境袭击风险陡增，单一国家的安全防御体系早已难以应对复合型安全威胁。更关键的是，美国长期奉行“双重标准”，频繁透支区域安全信誉，对盟友的安全承诺摇摆不定，让中东各国彻底认清“依附外部霸权无长久安全”的现实。在此背景下，麦加协议加速走出形式化困境，一系列实质性动作接连落地，彻底改写了协议的初始定位。

麦加协议由虚向实的最核心标志，是合作机制从模糊框架走向常态化、具象化实操。短短数月间，沙土巴三国打破分歧壁垒，构建起多层级沟通体系，防务高官常态化磋商、军事技术对接、联合防御推演逐步落地，彻底改变此前“只签约、不落地”的僵局。麦加协议已形成明确的协作逻辑，即以集体防御为抓手，统筹三方空域防御、跨境反恐、危机响应、情报共享等关键领域合作，将抽象的“共同防务”转化为可落地、可执行的安全举措。更具突破性的是，三国摆脱阵营对抗的固化思维，主动推动区域安全包容共治。从对抗博弈到对话共治，从依附域外到自主统筹，麦加协议的务实转型，彻底颠覆了外界对中东区域合作的固有认知。

麦加协议的实质性落地，正在深度重塑中东乃至欧亚大陆的地缘安全格局，其战略影响深远。

其一，逐步瓦解美国单边主导的中东安全体系。麦加协议的落地成型，打造出不依附于西方的本土安全合作平台，沙特阿拉伯、土耳其等传统美国盟友，开始跳出单一安全依附体系，通过区域抱团掌握安全主动权。这意味着美国在中东的“绝对话语权”大幅削弱，域外势力干预中东事务的抓手持续减少，中东安全格局正式进入“去霸权化”的多元变革期。

其二，深度重塑区域力量平衡，消解教派与国别分歧。中东长期因教派差异、历史矛盾、利益纷争陷入分裂对立，难以形成合力。麦加协议以共同安全利益为最大公约数，超越教派、国别、政体差异，让立场迥异的区域大国实现深度防务协作。这种“求同存异、共治共赢”的合作模式，正在逐步弥合中东内部裂痕，推动区域国家从“各自为战”转向“抱团自治”，为中东持久稳定奠定了基础。

其三，探索发展中国家地缘自主的全新范本。在全球地缘博弈加剧、霸权主义抬头的背景下，多数区域小国长期被动依附大国、夹缝求生。麦加协议由虚向实的蜕变证明，区域国家可以通过内部整合、自主协作，构建独立、包容、可持续的安全架构，无需依附单一域外势力。这一实践不仅为中东突破地缘困境提供了路径，也为全球南方国家争取战略自主、完善区域治理提供了借鉴。

尽管麦加协议由虚向实过程中仍面临多重现实挑战，但不可否认的是，中东战略自主的时代大势正在成型，麦加协议从一纸空文到实质落地，是中东地缘格局演变的必然结果，是区域国家摆脱外部依附、追求自主安全、实现共治共赢的主动选择。麦加协议最核心的价值是唤醒了中东的区域主体意识，构建了本土化的安全治理逻辑。站在全球格局变革的视角来看，麦加协议的蜕变，是世界多极化趋势在中东的生动落地。当霸权秩序日渐式微、阵营对抗不得人心，区域自主、包容共治、合作共赢已然成为全球地缘发展的主流。未来，随着机制持续完善、合作不断深化、伙伴逐步扩容，麦加协议将进一步夯实中东自主安全架构，推动中东走向稳定、自主、共赢的全新发展阶段，为全球地缘治理体系变革注入持久的中东力量。

（作者系国务院发展研究中心亚非发展研究所所长）

国际时评

欧盟不当域外管辖的实质是单边垄断

□谢 军

近期，欧盟对京东进行调查，肆意向中国实体跨境索取广泛的、非必要的中国境内信息。对此，我国认定欧盟利用其《关于扭曲欧盟内部市场的外国补贴条例》（以下简称《外国补贴条例》，英文简称“FSR”）对中国实体采取的相关跨境调查做法，构成不当域外管辖。这一事件随即引发国际社会广泛关注。欧盟《外国补贴条例》于2023年7月12日开始正式实施，是其自行创设的单边工具。近年来，欧盟频繁依托该条例在全球发起“外国补贴”相关调查，实施不当域外管辖行为，凸显出鲜明的贸易保护与制度排他属性。

欧盟依托《外国补贴条例》实施具有单边垄断性质的不当域外管辖，体现在以下方面。其一，管辖依据泛化，涉嫌违反国际法与国际关系基本准则。欧盟《外国补贴条例》条文标准模糊、自由裁量空间极大，肆意将各类境外经营主体与经营信息纳入审查范围，严重违背国际法基本准则。其二，调查行为肆意越界，非法干涉他国主权。欧盟跨境调查不仅针对涉案企业，还强行突破地域管辖边界，强制要求他国境内金融机构、合作企业等第三方主体配合取证，肆意索取大量与案件无关的境内经营、金融、政务信息，严重违背属地管辖基本原则与跨境监管权责边界。2025年2月，在对韩国水力原子力公司（KHNP）参与捷克杜科瓦尼核电站招标项目的审查过程中，欧委会以外国补贴审查为由，要求KHNP提交韩国本土政府财政支持材料。该审查虽最终未启动深入调查即结案，却把欧盟竞争规则的效力拓展至第三国政府的财政举措。这一做法暴露了FSR的问题，即它借调查越界获取域外信息、介入他国事务，使欧委会的审查权触及他国主权管辖下的经济治理与市场秩序。其三，适用场景持续扩容，形成常态化制度性打压。欧盟域外管辖已从传统技术安全、高端安防等特殊领域，全面延伸至普通商业竞争赛道。2025年7月，阿联酋阿布扎比国家石油公司（ADNOC）收购德国科思创（Covestro）股权项目，该交易已通过欧盟合并审查，欧委会仍依《外国补贴条例》启动深度调查，认定无限额国家担保、增资承诺及优惠税收构成补贴，最终附条件批准交易，要求取消国家担保并强制许可科思创专利。此举意味着FSR的审查范围已从传统国家安全敏感领域扩展至化工等普通商业并购，域外监管工具的适用场景持续扩容，常态化、制度性打压的趋向日益显现。

欧盟频繁进行不当域外管辖主要有以下动因。第一，维护自身经贸优势，实施隐性贸易保护。随着新兴经济体企业在数字经济、跨境商贸、高端制造等领域全球竞争力持续攀升，欧盟传统产业面临显著竞争压力。为稳固区域经贸垄断优势，欧盟企图依托宽泛的域外监管规则构筑隐形非关税贸易壁垒，大幅抬高欧盟境外企业进入欧盟市场的准入门槛，实现对域外优质竞争者的制度性压制。第二，争夺规则主导权，重塑多边贸易格局。以世界贸易组织（WTO）为核心的多边贸易体系，对各类贸易保护举措形成一定程度的约束，与欧盟近年来日益增长的谋求排他性竞争优势的目标存在冲突。欧盟规避多边争端解决与监管框架，依托自身经济体量构建单边审查体系，以自身监管标准替代国际通行规则，意在抢占全球经贸规则制定话语权。第三，服务地缘战略目标，遏制全球新兴产业崛起。欧盟域外监管持续升级，不断脱离单纯经贸范畴，带有鲜明战略指向。其依托常态化、全行业域外管辖压制新兴经济体出海企业，阻碍跨国企业全球布局、抑制新兴产业发展，本质是服务于地缘政治制衡的目标。

欧盟不当域外管辖的常态化扩张与制度化滥用，对国际经贸秩序、各国经济主权与多边合作体系造成巨大挑战。首先，单边监管壁垒严重扭曲全球公平竞争格局。欧盟依托《外国补贴条例》模糊的判定标准与极大的自由裁量权，对欧盟境外企业实施无差别严苛审查，人为抬高国际企业入驻欧盟市场的制度门槛。其挑战公平竞争原则，以监管之名设置隐性非关税贸易壁垒，压制境外竞争者，打破了国际贸易平等互利、自由竞争的准则，严重破坏全球市场竞争生态。其次，越界监管行为挑战各国经济主权与国际法根基。欧盟在跨境调查中，肆意突破地域监管边界，强制要求多国境内第三方机构配合取证，肆意调取他国境内经营、金融、政务等属地核心信息，无端介入他国内部经济事务。这种跨界权力越位行为，无视国际法基本原则，践踏各国经济主权，严重削弱了国际法的约束力与权威性。最后，制度化单边打压加剧全球经贸体系碎片化。当前以WTO为核心的多边贸易体系，倡导互利共赢的治理理念，是全球经贸稳定运行的基石。欧盟刻意绕开多边规则体系，依托自身经济体量构建单边域外审查机制，以自身监管标准替代国际通行规则，强行输出自身制度范式。这种排他性、霸权式的监管模式，加剧了全球经贸规则隔阂与对立，扰乱了全球产业链、供应链的稳定运转，阻碍了全球经济治理朝着公平化、规范化、一体化的方向推进，给全球跨境贸易与投资合作带来极大不确定性。

针对欧盟常态化、制度化的不当域外管辖行为，全球各经济体应构建长效应对机制。首先，完善涉外法治体系，夯实制度保障根基。可借鉴多国成熟法治实践，补齐涉外监管阻断制度短板，搭建集风险识别、研判评估、合规反制于一体的工作机制。针对经贸、金融、科技、资源投资等不同领域的域外打压场景，细化不当管辖认定标准与分级应对规则，明晰市场主体权责边界，畅通企业合规救济渠道，构建兼具规范性与实操性的涉外法治体系。其次，规范法治实操路径，构建精准阻断机制。应对域外监管干预始终秉持依法合规、客观审慎的原则，以国际法等相关法律法规为依据，甄别各类越界监管行为。坚持针对不符合法理的不当域外管辖实施合规阻断，杜绝随意化处置，以专业、规范的法治实践维护跨境经贸治理公信力。最后，参与全球规则治理，重塑公平贸易秩序。各经济体应积极依托联合国、WTO、金砖国家（BRICS）、上海合作组织等多边合作平台，凝聚抵制单边长臂管辖、反对贸易保护的国际共识，倡导公平公正、互利共赢的治理理念，共同捍卫以国际法为基础的全球经贸秩序，推动全球经济治理体系朝着更加开放、包容、公正、合理的方向发展。

国别观察