

AI 治理有望为中美关系注入积极内容

让人工智能成为能够管理竞争、增加确定性的新合作领域，可能正是当前中美关系最值得认真探索的现实选择。

2026 年 9 月的中美元首峰会，已经在全球瞩目中结束。峰会达成的八点共识中，双方同意建立人工智能对话，交流人工智能相关风险和惠益。这个共识的达成，意味着人工智能治理，正在从科技竞争议题进入中美建设性战略稳定关系相关的议程之中。

这并非偶然。过去一段时间，美国围绕人工智能安全、前沿模型风险和军事应用的讨论明显升温。美国政府一方面持续推动 AI 进入情报、国防和国家安全体系，另一方面又开始更加频繁地讨论模型失控、网络攻击、生物风险、军事误判以及人类控制等问题。美国头部 AI 企业也密集发布安全报告，推动前沿模型评估、重大安全事件报告和高风险能力治理。

同时，据 CNN 援引多名知情人士报道，今年春季，美国军方曾根据一份 AI 辅助生成的情报报告，判断一艘在中东航行的中国船只可能运输与核武器项目有关的部件。随后，美军开始准备武装登临，军机已经升空。行动进入最后阶段后，专业人员重新检查底层材料，发现 AI 聊天机器人错误识别了船上货物，行动最终被叫停。这个案例已经足够提出一个严肃的问题：如果一次 AI 幻觉进入军事情报流程之后没有被及时发现，会发生什么？

在这样的背景下，人工智能治理被纳入与中美建设性战略稳定相关的议程，就变得顺理成章。过去讨论大国战略稳定，重点通常包括核误判、预警系统故障、军事力量意



沈逸

外接触、指挥通信失灵等。进入 AI 时代之后，急需构建人工智能事件的沟通渠道。

AI 的优势在于速度。同样的优势也可能成为风险来源。机器能够在几秒钟内完成过去需要分析人员数小时甚至数天完成的工作，也就意味着错误同样可以在几秒钟内进入后续决策流程。当情报搜集、目标识别、威胁判断和行动规划越来越多地引入 AI，人类进行交叉核验、政治判断和危机沟通的时间可能进一步缩短。因此，中美开展 AI 安全合作具有现实的危机管理价值，重点可以放在风险控制和防止技术故障升级为安全危机，并不需要双方共享核心算法、军事模型或敏感数据。这构成了 AI 治理可能成为中美峰会一个建设性议题的现实基础。

先进模型、算力、芯片、产业生态都已经成为两国科技竞争的重要组成部分。双方对全球 AI 治理的理念也存在差异。美国部分政策力量仍然习惯从技术领先和地缘竞争理解 AI 问题，中国则长期主张发展与安全并重，坚持人工智能造福全人类，推动形成广泛参与、公平合理的全球治理体系。如果双方能够从 AI 安全和战略稳定入手，围绕人类控制、危机沟通和高风险应用形成一些具体、可核验、能够逐步推进的安排，其意义将远超预期。中美需要竞争规则，也需要风险护栏。让人工智能成为能够管理竞争、增加确定性的新合作领域，可能正是当前中美关系最值得认真探索的现实选择。

中美开展 AI 安全合作具有现实的危机管理价值。

✉ [发自上海]