

# 人工智能 不能失控

当机器开始思考、  
算法参与决策、技术挑  
战伦理、鸿沟不断拉大，  
我们应该如何应对

当机器开始思考，人类如何与之相处？当算法参与决策，安全如何保障？当技术挑战伦理，治理如何跟上？当鸿沟不断拉大，普惠如何实现？

7月17日，国家主席习近平在2026世界人工智能大会暨人工智能全球治理高级别会议开幕式上提出了四个“时代之问”，引发高度关注。

AI发展是大势所趋，但同时算法欺诈、隐私泄露、深度伪造等问题，也带来不容忽视的安全风险与伦理挑战。我们应该如何应对？

## 发展越猛 安全的弦越不能松

先聊智能。人工智能正深刻重塑着人类生产生活方式。而更关键的变化在于，伴随着具身智能的迭代，AI正从数字世界走向现实世界，从“能思考”走向“能干活”。如今，已经有具备灵巧手、掌握专项技能的具身机器人走上劳动岗位。

再说算法。过去，算法能推荐一条新闻、带火一段视频；而如今，它已经融入医疗辅助诊断、金融风险研判、城市交通调度等领域，甚至参与科研。在不远的将来，它还可能与更多智能技术结合，实现“上天下海”。

总览行业。截至2025年底，我国人工智能企业已超过6000家，从基础底座、模型框架到行业应用，形成了一套完整的产业体系；智能经济核心产业规模突破1.2万亿元；一批国产大模型凭借开源战略引领全球生态，把曾经只被少数企业掌握的尖端科技，变成千行百业可调用的普惠工具。

发展的势头越猛，安全的弦越不能松。不发展，就抓不住这轮科技革命的战略主动；不安全，那即便跑得再快也迟早会撞上“南墙”。

## 争议核心 谁该为创作负责

这场大会，恰逢其时。因为人工智能的伦理风险，正试探着现行规则的底线，这逼着我们必须坐下来，把规则重新梳理明白。

拿这些天的两场争议来说，呼和浩特一场摄影比赛，一等奖作品被确认系AI生成，画面里的环卫工人“共用一张脸”，衣服上的文字成了乱码；东莞一场书法征集，二等奖作品也因疑似AI生成被撤下。

相关争议的核心，不是AI能不能创作，而是谁该为创作负责。人工智能本就是人类造出来的新工具，一概拒绝，既不现实，也无必要。真正要警惕的，是有人借技术之便，模糊创作的边界，以机器生成的东西冒充人的劳动。

此外，大会开幕前几天，中国知网也发布声明，把DeepSeek、Gemini等人工智能列为作者的论文，一律下架，理由讲得干脆：AI不具备民事主体资格，既无民事权利能力，亦无民事行为能力；一旦学术出了问题，它无法承担核查、整改、追责的义务。把AI推上署名位，等于抽掉了责任主体。

这些看得见的风波，不过是冰山一角。更隐蔽的风险，来自大模型的“黑箱”特性。

与传统网络安全相比，AI领域的攻防呈现鲜明的非对称性：模型幻觉、算法偏见、内在不可解释性等内生风险，与技术落地过程中滋生的开源滥用、深度伪造、情感诱导等衍生风险，相互交织、彼此放大。它们触点多、传导快、跨界性、关联性、穿透性极强——这不是局限于某些行业的风险，而是整个社会未来要面临的考题。

## 如何治理 始终由人类控制

面对技术洪流外溢所产生的风险与挑战，治理该如何跟上？

中国给出了清晰答案：要高度重视人工智能引发的各类内生和衍生风险，推动构建法律法规、技术监测、风险预警、应急响应体系，筑牢安全底线，防范滥用恶用，确保人工智能始终处于人类控制之下。

这一方案的逻辑清晰而严密。在对象这侧，既指向模型幻觉、算法偏见、大模型内在不可解释等技术内生风险，也指向开源生态被滥用、软件出现漏洞、深度伪造污染信息环境等技术外化于社会的衍生风险。

在手段这边，法律法规是“骨架”，划定权责边界——从《生成式人工智能服务管理暂行办法》到《人工智能生成合成内容标识办法》，算法备案、大模型安全评估、内容标识正逐项落地；技术监测是“耳目”，第三方审计与认证机构持续检测模型的对抗安全、后门安全与可解释性；风险预警是“哨兵”，定期报送风险评估报告，把隐患拦在爆发之前；应急响应是“底线”，一旦出事，第一时间干预、止损、追责。四者环环相扣，才谈得上安全可控。

“确保人工智能始终处于人类控制之下”，则是AI治理的长期目标。无论人机如何协作，在关键的决策节点上，价值判断的最终一票，必须攥在人的手里。

值得一提的是，近年来，个别国家把人工智能当成国际竞争的战略筹码，热衷出口管制、构筑“小院高墙”，以安全之名行遏制之实。可安全从来不是零和的独木桥，在深度互联的智能时代，没有哪个国家能筑起一道墙，把风险统统挡在门外。真正的安全，是共同的安全；真正的底线，要靠各方一起来守。

这也正是中国方案的分量所在。本届大会上，29个创始成员国签署协定，世界人工智能合作组织落户上海；《人工智能合作发展行动计划》《国际人工智能伦理治理行动计划》相继发布。一年前的构想，如今化作落地成果。

规则越透明，创新越敢前行；边界越清楚，探索越有底气。清晰的治理方向是给创新开路的最好方式。

机器会思考，是这个时代最激动人心的进步。但人类更要清醒，因为需要思考的，不只是机器，还有我们自己：技术究竟为了谁？创新最终服务谁？未来到底属于谁？

答案或许并不复杂：无论走多远，人工智能的指针，必须始终指向人类福祉；无论变得多强，它必须是人类所能掌握的工具。

上游财经综合重庆日报